

PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau



AM

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 7 : C12N 15/12, C07K 14/47, 16/18, G01N 33/566, C12Q 1/68, C12N 15/11, 15/62, A01K 67/027, A61K 38/00		A2	(11) International Publication Number: WO 00/58473 (43) International Publication Date: 5 October 2000 (05.10.00)
(21) International Application Number: PCT/US00/08621 (22) International Filing Date: 31 March 2000 (31.03.00) (30) Priority Data: 60/127,607 31 March 1999 (31.03.99) US 60/127,636 2 April 1999 (02.04.99) US 60/127,728 5 April 1999 (05.04.99) US 09/540,763 30 March 2000 (30.03.00) US (63) Related by Continuation (CON) or Continuation-in-Part (CIP) to Earlier Applications US 60/127,607 (CIP) Filed on 31 March 1999 (31.03.99) US 60/127,636 (CIP) Filed on 2 April 1999 (02.04.99) US 60/127,728 (CIP) Filed on 5 April 1999 (05.04.99) US 09/540,763 (CIP) Filed on 30 March 2000 (30.03.00) (71) Applicant (for all designated States except US): CURAGEN CORPORATION [US/US]; 555 Long Wharf Drive, 11th Floor, New Haven, CT 06511 (US).		(72) Inventors; and (75) Inventors/Applicants (for US only): SHIMKETS, Richard, A. [US/US]; 191 Leete Street, West Haven, CT 06516 (US). LEACH, Martin [GB/US]; 884 School Street, Webster, MA 01570 (US). (74) Agent: ELRIFI, Ivor, R.; Mintz, Levin, Cohn, Ferris, Glovsky and Popeo, P.C., One Financial Center, Boston, MA 02111 (US). (81) Designated States: AE, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CR, CU, CZ, DE, DK, DM, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).	
		Published Without international search report and to be republished upon receipt of that report.	
(54) Title: NUCLEIC ACIDS INCLUDING OPEN READING FRAMES ENCODING POLYPEPTIDES; "ORFX"			
(57) Abstract <p>The present invention provides open reading frames ORFX, encoding isolated polypeptides, as well as polynucleotides encoding ORFX and antibodies that immunospecifically bind to ORFX or any derivative, variant, mutant, or fragment of the ORFX polypeptides, polynucleotides or antibodies. The invention additionally provides methods in which the ORFX polypeptide, polynucleotide and antibody are used in detection and treatment of a broad range of pathological states, as well as to other uses.</p>			

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece	ML	Mali	TR	Turkey
BG	Bulgaria	HU	Hungary	MN	Mongolia	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MR	Mauritania	UA	Ukraine
BR	Brazil	IL	Israel	MW	Malawi	UG	Uganda
BY	Belarus	IS	Iceland	MX	Mexico	US	United States of America
CA	Canada	IT	Italy	NE	Niger	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NL	Netherlands	VN	Viet Nam
CG	Congo	KE	Kenya	NO	Norway	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NZ	New Zealand	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	PL	Poland		
CM	Cameroon	KR	Republic of Korea	PT	Portugal		
CN	China	KZ	Kazakstan	RO	Romania		
CU	Cuba	LC	Saint Lucia	RU	Russian Federation		
CZ	Czech Republic	LJ	Liechtenstein	SD	Sudan		
DE	Germany	LK	Sri Lanka	SE	Sweden		
DK	Denmark	LR	Liberia	SG	Singapore		
EE	Estonia						

NOVEL POLYNUCLEOTIDES AND POLYPEPTIDES ENCODED THEREBY

5

BACKGROUND OF THE INVENTION

The invention relates generally to nucleic acids and polypeptides encoded thereby, and methods of using these nucleic acids and polypeptides.

10

SUMMARY OF THE INVENTION

The invention is based in part on the discovery of nucleic acids that include open reading frames encoding novel polypeptides, and on the polypeptides encoded thereby. The nucleic acids and polypeptides are collectively referred to herein as "ORFX".

Accordingly, in one aspect, the invention provides an isolated nucleic acid molecule (SEQ ID NO:2 n -1, wherein n is an integer between 1-3161), that encodes novel polypeptide, or a fragment, homolog, analog or derivative thereof. The nucleic acid can include, *e.g.*, a nucleic acid sequence encoding a polypeptide at least 85% identical to a polypeptide comprising the amino acid sequences of SEQ ID NO:2 n , wherein n is an integer between 1-3161. The nucleic acid can be, *e.g.*, a genomic DNA fragment, or a cDNA molecule.

Also included in the invention is a vector containing one or more of the nucleic acids described herein, and a cell containing the vectors or nucleic acids described herein.

The invention is also directed to host cells transformed with a recombinant expression vector comprising any of the nucleic acid molecules described above.

In another aspect, the invention includes a pharmaceutical composition that includes an ORFX nucleic acid and a pharmaceutically acceptable carrier or diluent.

In a further aspect, the invention includes a substantially purified ORF polypeptide, *e.g.*, any of the ORFX polypeptides encoded by an ORFX nucleic acid, and fragments, homologs, analogs, and derivatives thereof. The invention also includes a pharmaceutical composition that includes a ORFX polypeptide and a pharmaceutically acceptable carrier or diluent.

5 In a still a further aspect, the invention provides an antibody that binds specifically to an ORFX polypeptide. The antibody can be, *e.g.*, a monoclonal or polyclonal antibody, and fragments, homologs, analogs, and derivatives thereof. The invention also includes a pharmaceutical composition including ORFX antibody and a pharmaceutically acceptable carrier or diluent. The invention is also directed to isolated antibodies that bind to an epitope on a
10 polypeptide encoded by any of the nucleic acid molecules described above.

The invention also includes kits comprising any of the pharmaceutical compositions described above.

The invention further provides a method for producing an ORFX polypeptide by providing a cell containing a ORFX nucleic acid, *e.g.*, a vector that includes a ORFX nucleic
15 acid, and culturing the cell under conditions sufficient to express the ORFX polypeptide encoded by the nucleic acid. The expressed ORFX polypeptide is then recovered from the cell. Preferably, the cell produces little or no endogenous ORFX polypeptide. The cell can be, *e.g.*, a prokaryotic cell or eukaryotic cell.

The invention is also directed to methods of identifying an ORFX polypeptide or nucleic
20 acids in a sample by contacting the sample with a compound that specifically binds to the polypeptide or nucleic acid, and detecting complex formation, if present.

The invention further provides methods of identifying a compound that modulates the activity of a ORFX polypeptide by contacting ORFX polypeptide with a compound and determining whether the ORFX polypeptide activity is modified.

25 The invention is also directed to compounds that modulate ORFX polypeptide activity identified by contacting a ORFX polypeptide with the compound and determining whether the compound modifies activity of the ORFX polypeptide, binds to the ORFX polypeptide, or binds to a nucleic acid molecule encoding a ORFX polypeptide.

30 In a another aspect, the invention provides a method of determining the presence of or predisposition of an ORFX-associated disorder in a subject. The method includes providing a sample from the subject and measuring the amount of ORFX polypeptide in the subject sample.

The amount of ORFX polypeptide in the subject sample is then compared to the amount of ORFX polypeptide in a control sample. An alteration in the amount of ORFX polypeptide in the subject protein sample relative to the amount of ORFX polypeptide in the control protein sample indicates the subject has a tissue proliferation-associated condition. A control sample is
5 preferably taken from a matched individual, *i.e.*, an individual of similar age, sex, or other general condition but who is not suspected of having a tissue proliferation-associated condition. Alternatively, the control sample may be taken from the subject at a time when the subject is not suspected of having a tissue proliferation-associated disorder. In some embodiments, the ORFX is detected using a ORFX antibody.

10 In a further aspect, the invention provides a method of determining the presence of or predisposition of an ORFX-associated disorder in a subject. The method includes providing a nucleic acid sample, *e.g.*, RNA or DNA, or both, from the subject and measuring the amount of the ORFX nucleic acid in the subject nucleic acid sample. The amount of ORFX nucleic acid sample in the subject nucleic acid is then compared to the amount of an ORFX nucleic acid in a
15 control sample. An alteration in the amount of ORFX nucleic acid in the sample relative to the amount of ORFX in the control sample indicates the subject has a tissue proliferation-associated disorder.

In a still further aspect, the invention provides method of treating or preventing or delaying a ORFX-associated disorder. The method includes administering to a subject in which
20 such treatment or prevention or delay is desired a ORFX nucleic acid, a ORFX polypeptide, or an ORFX antibody in an amount sufficient to treat, prevent, or delay a tissue proliferation-associated disorder in the subject.

Unless otherwise defined, all technical and scientific terms used herein have the same meaning as commonly understood by one of ordinary skill in the art to which this invention
25 belongs. Although methods and materials similar or equivalent to those described herein can be used in the practice or testing of the present invention, suitable methods and materials are described below. All publications, patent applications, patents, and other references mentioned herein are incorporated by reference in their entirety. In the case of conflict, the present specification, including definitions, will control. In addition, the materials, methods, and
30 examples are illustrative only and not intended to be limiting.

Other features and advantages of the invention will be apparent from the following detailed description and claims.

DETAILED DESCRIPTION OF THE INVENTION

The invention provides novel polypeptides and nucleotides encoded thereby. The
5 polynucleotides and their encoded polypeptides can be grouped according to the functions played by their gene products. Such functions include, structural proteins, proteins from which associated with metabolic pathways fatty acid metabolism, glycolysis, intermediary metabolism, calcium metabolism, proteases, and amino acid metabolism, etc.

Included in the invention are 3161 novel nucleic acid sequences and their encoded
10 polypeptides. The sequences are collectively referred to as "ORFX nucleic acids" or ORFX polynucleotides" and the corresponding encoded polypeptide is referred to as a "ORFX polypeptide" or ORFX protein". For example, an ORFX nucleic acid according to the invention is a nucleic acid including an ORF1 nucleic acid, and an ORF polypeptide according to the invention is a polypeptide that includes the amino acid sequence of an ORF1 polypeptide.
15 Unless indicated otherwise, "ORFX" is meant to refer to any of the ORF1-3161 sequences disclosed herein.

Table 1 provides a summary of the ORFX nucleic acids and their encoded polypeptides are summarized in Table 1. Nucleic acid sequences and polypeptide sequences for ORFX nucleic acids according to the invention is provided in the section of the specification entitled
20 "Disclosed Sequences of ORFX Nucleic Acid and Polypeptide Sequences."

Column 1 of Table 1, entitled "ORF #", denotes an ORF number assigned to a nucleic acid containing an open reading frame according to the invention.

Column 2 of Table 1, entitled "Internal Identification number (Nucleic Acid Sequence Identification Number, Polypeptide Sequence Identification Number), provides an internal
25 identification number for the indicated ORF, along with sequence identification numbers (SEQ ID NOs.) corresponding to the indicated ORF. In general, for an ORF n according to the invention (wherein n is any integer from 1 to 3161), a nucleic acid corresponding to the ORF is SEQ ID NO:2 n -1, and an amino acid sequence encoded by the ORF is SEQ ID NO:2 n . For example, a nucleic acid sequence corresponding to an ORF1 nucleic acid is SEQ ID NO:1, and a
30 polypeptide sequence corresponding to an ORF1 polypeptide is SEQ ID NO:2. Similarly, a

nucleic acid sequence corresponding to an ORF4 nucleic acid is SEQ ID NO:7, and a polypeptide sequence corresponding to an ORF4 polypeptide is SEQ ID NO:8; a nucleic acid sequence corresponding to an ORF198 nucleic acid sequence is SEQ ID NO:395, and a polypeptide sequence corresponding to an ORF198 polypeptide is SEQ ID NO:396. Nucleic acid sequences and polypeptide sequences for ORFX nucleic acids according to the invention are provided in the section of the specification entitled "Disclosed Sequences of ORFX Nucleic Acid and Polypeptide Sequences."

Column 2 of Table 1, entitled "Protein Similarity", lists previously described proteins that are related to polypeptides encoded by the ORFs. Genbank identifiers for the previously described proteins are provided. These can be retrieved from <http://www.ncbi.nlm.nih.gov/>.

To determine similarity to previously described proteins, polypeptides encoded by ORFX DNA sequences were tested using the Framesearch Algorithm against a nonredundant version of the GenPept Database from NCBI/Genbank. DNA sequences that had a score of '90' or above (Framesearch algorithm score, Edelman et. al. GCG Genetics) to a known protein were selected. Open reading frames were extended beyond the region of the protein matched using standard DNA translation and codon tables. Novel proteins that lacked a protein match were translated against the standard genetic codons and proteins with an ORF at least 80 amino acids and containing a Methionine start are included in the Table.

Column 3 of Table 3, entitled "Protein Domains", lists previously described protein domains, designated by pfam entries, that are present in polypeptides encoded by the ORFs. Also included in column 3 are proteins in which these domains are present. The pfam entries can be retrieved from <http://pfam.wustl.edu/>. DNA sequences were translated in all six frames and tested using the Hmmer Algorithm against the Pfam Database (References to the algorithm and Pfam database can be found at <http://pfam.wustl.edu/>). Translated DNA sequences that matched a protein domain entry in the Pfam database AND had a score of '7.5' were selected.

Column 4 of Table 3, entitled "Protein Classification", lists the type of classification assigned for the protein, based on its homology. Examples of proteins in the classification include the following proteins:

Amylases

Amylase is responsible for endohydrolysis of 1,4-alpha-glucosidic linkages in oligosaccharides and polysaccharides. Variations in amylase gene may be indicative of delayed maturation and of various amylase producing neoplasms and carcinomas.

5 **Amyloid**

The serum amyloid A (SAA) proteins comprise a family of vertebrate proteins that associate predominantly with high density lipoproteins (HDL). The synthesis of certain members of the family is greatly increased in inflammation. Prolonged elevation of plasma SAA levels, as in chronic inflammation, 15 results in a pathological condition, called amyloidosis, which affects the liver, kidney and spleen and which is characterized by the highly insoluble accumulation of SAA in these tissues. Amyloid selectively inhibits insulin-stimulated glucose utilization and glycogen deposition in muscle, while not affecting adipocyte glucose metabolism. Deposition of fibrillar amyloid proteins intraneuronally, as neurofibrillary tangles, extracellularly, as plaques and in blood vessels, is characteristic of both Alzheimer's disease and aged Down's syndrome. 15 Amyloid deposition is also associated with type II diabetes mellitus.

Angiopoeitin

Members of the angiopoeitin/fibrinogen family have been shown to stimulate the generation of new blood vessels, inhibit the generation of new blood vessels, and perform several roles in blood clotting. This generation of new blood vessels, called angiogenesis, is also an 20 essential step in tumor growth in order for the tumor to get the blood supply it needs to expand. Variation in these genes may be predictive of any form of heart disease, numerous blood clotting disorders, stroke, hypertension and predisposition to tumor formation and metastasis. In particular, these variants may be predictive of the response to various antihypertensive drugs and chemotherapeutic and anti-tumor agents.

25 **Apoptosis-related proteins**

Active cell suicide (apoptosis) is induced by events such as growth factor withdrawal and toxins. It is controlled by regulators, which have either an inhibitory effect on programmed cell

death (anti-apoptotic) or block the protective effect of inhibitors (pro-apoptotic). Many viruses have found a way of countering defensive apoptosis by encoding their own anti-apoptosis genes preventing their target-cells from dying too soon. Variants of apoptosis related genes may be useful in formulation of anti-aging drugs.

5 **Cadherin, Cyclin, Polymerase, Oncogenes, Histones, Kinases**

Members of the cell division/cell cycle pathways such as cyclins, many transcription factors and kinases, DNA polymerases, histones, helicases and other oncogenes play a critical role in carcinogenesis where the uncontrolled proliferation of cells leads to tumor formation and eventually metastasis. Variation in these genes may be predictive of predisposition to any form
10 of cancer, from increased risk of tumor formation to increased rate of metastasis. In particular, these variants may be predictive of the response to various chemotherapeutic and anti-tumor agents.

Colony-stimulating factor-related proteins

Granulocyte/macrophage colony-stimulating factors are cytokines that act in
15 hematopoiesis by controlling the production, differentiation, and function of 2 related white cell populations of the blood, the granulocytes and the monocytes-macrophages.

Complement-related proteins

Complement proteins are immune associated cytotoxic agents, acting in a chain reaction to exterminate target cells to that were opsonized (primed) with antibodies, by forming a
20 membrane attack complex (MAC). The mechanism of killing is by opening pores in the target cell membrane. Variations in 20 complement genes or their inhibitors are associated with many autoimmune disorders. Modified serum levels of complement products cause edemas of various tissues, lupus (SLE), vasculitis, glomerulonephritis, renal failure, hemolytic anemia, thrombocytopenia, and arthritis. They interfere with mechanisms of ADCC (antibody dependent
25 cell cytotoxicity), severely impair immune competence and reduce phagocytic ability. Variants of complement genes may also be indicative of type I diabetes mellitus, meningitis neurological disorders such as nemaline myopathy, neonatal hypotonia, muscular disorders such as congenital myopathy and other diseases.

Cytochrome

The respiratory chain is a key biochemical pathway which is essential to all aerobic cells. There are five different cytochromes involved in the chain. These are heme bound proteins which serve as electron carriers. Modifications in these genes may be predictive of ataxia
5 areflexia, dementia and myopathic and neuropathic changes in muscles. Also, association with various types of solid tumors.

Kinesins

Kinesins are tubulin molecular motors that function to transport organelles within cells and to move chromosomes along microtubules during cell division. Modifications of these genes
10 may be indicative of neurological disorders such as Pick disease of the brain, tuberous sclerosis.

Cytokines, Interferon, Interleukin

Members of the cytokine families are known for their potent ability to stimulate cell growth and division even at low concentrations. Cytokines such as erythropoietin are cell-specific in their growth stimulation; erythropoietin is useful for the stimulation of the
15 proliferation of erythroblasts. Variants in cytokines may be predictive for a wide variety of diseases, including cancer predisposition.

G-protein coupled receptors

G-protein coupled receptors (also called R7G) are an extensive group of hormones, neurotransmitters, odorants and light receptors which transduce extracellular signals by
20 interaction with guanine nucleotide-binding (G) proteins. Alterations in genes coding for G-coupled proteins may be involved in and indicative of a vast number of physiological conditions. These include blood pressure regulation, renal dysfunctions, male infertility, dopamine associated cognitive, emotional, and endocrine functions, hypercalcemia, chondrodysplasia and osteoporosis, pseudohypoparathyroidism, growth retardation and
25 dwarfism.

Thioesterases

Eukaryotic thiol proteases are a family of proteolytic enzymes which contain an active site cysteine. Catalysis proceeds through a thioester intermediate and is facilitated by a nearby histidine side chain; an asparagine completes the essential catalytic triad. Variants of thioester associated genes may be predictive of neuronal disorders and mental illnesses such as Ceroid Lipoffiscinosis, Neuronal 1, Infantile, Santavuori disease and more.

The key to the molecule type is as follows:

10	Abbrev:	Title:
	amylase	amylase protein
	amylaseinhib	amylase inhibitor
	amyloid	amyloid protein
15	apoptosis	apoptosis associated protein
	apoptosisinhib	apoptosis inhibitors
	apoptosisrecep	apoptosis receptors
	ATPase_associated	ATPase associated protein
	biotindep	biotin dependent enzyme/protein
20	cadherin	cadherin protein
	calcium_channel	calcium channel protein
	carboxylase	carboxylase protein
	cathepsin	cathepsin/carboxypeptidases
	cathepsininhib	cathepsin/carboxypeptidase inhibitor
25	chloride_channel	chloride channel protein
	collagen	collagen
	complement	complement protein
	complementrecept	complement receptor protein
	complementinhib	complement inhibitor
30	csf	colony stimulating factor
	csfrecept	colony stimulating factor receptor
	cyclin	cyclin protein
	cyto450	cytochrome p450 protein
	cytochrome	cytochrome related protein
35	deaminase	deaminase
	dehydrogenase	dehydrogenase
	desaturase	desaturase
	dna_rna_bind	DNA/RNA binding protein/factor
	dna_rna_inhib	DNA/RNA binding protein/factor inhibitor
40	dynein	dynein

	elastase	elastase
	elastaseinhib	elastase inhibitor
	eph	EPH family of tyrosine kinases
	esterase	esterase
5	esteraseinhib	esterase inhibitor
	fgf	fibroblast growth factor
	fgfreceptor	fibroblast growth factor receptor
	gaba	GABA receptor
	glucoamylase	glucoamylase
10	glucoronidase	glucoronidase
	glycoprotein	glycoprotein
	Guanylyl	guanylylate cyclase
	helicase	helicase
	histone	histone
15	HOM	homologous
	homeobox	homeobox protein
	hydrolase	hydrolase
	hydroxysteroid	hydroxysteroid associated protein
	hypoxanthine	hypoxanthine associated protein
20	immunoglob	immunoglobulin
	immunoglobrecept	immunoglobulin receptor
	interferon	interferon
	interleukin	interleukin
	interleukinrecept	interleukin receptor
25	isomerase	isomerase
	isomeraseinhibitor	isomerase inhibitor
	isomerasereceptor	isomerase receptor
	kinase	kinase
	kinaseinhibitor	kinase inhibitor
30	kinasereceptor	kinase receptor
	kinesin	kinesin
	laminin	laminin associated protein
	lipase	lipase
	metallothionein	metallothionein
35	MHC	major histocompatibility complex
	misc_channel	miscellaneous channel
	ngf	nerve growth factor
	nuci_recpt	nuclear receptor
	nuclease	nuclease
40	oncogene	oncogene associated protein
	oxidase	oxidase
	oxygenase	oxygenase
	peptidase	peptidase
	peroxidase	peroxidase
45	phosphatase	phosphatase
	phosphataseinhib	phosphatase inhibitor

	phosphorylase PIR	phosphorylase PIR DATABASE (release 56, 29-OCT-1998)
5	polymerase potassium_channel prostaglandin protease proteaseinhib reductase	polymerase potassium channel protein prostaglandin protease protease inhibitor reductase
10	ribosomalprot RTR	ribosomal associated protein EMBLDATABASE translated entries not to be incorporated into SWISS-PROT (20-JUL-1998)
15	SIM SPTR	similar EMBL DATABASE translated entries to be incorporated into SWISS-PROT (20-JUL-1998)
20	struct sulfotransferase SWP	structural associated protein sulfotransferase SWISS-PROT DATABASE (release 18-OCT-1998)
25	SWPN synthase tgf tgfreceptor thioesterase thiolase tm7	SWISS-PROT Update (release 11-NOV-98) synthase transforming growth factor transforming growth factor receptor thioesterase thiolase
30	tnf traffic tnfreceptor TRN	seven transmembrane domain G-protein coupled receptor necrosis factor receptor tumor necrosis factor tumor trafficking associated protein EMBL DATABASE translated entries update (20-JUL-1998)
35	transcriptfactor transferase transport tubulin ubiquitin unclassified	transcription factor transferase transport protein tubulin ubiquitin
40	water channel	Protein not categorized into one of the aforementioned protein families water channel protein

Column 5 of Table 1, entitled, "Cells or Tissues in Which Gene is Expressed", denotes tissues, represented by five digit numbers, in which RNA homologous to the ORF nucleic acid sequences is present. Tissues or cells corresponding to the numbers are provided in Table 2.

ORFX nucleic acids, and their encoded polypeptides, according to the invention are
5 useful in a variety of applications and contexts. For example, various ORFX nucleic acids and polypeptides according to the invention are useful, *inter alia*, as novel members of the protein families indicated in Table 1, and/or according to the presence of domains and sequence relatedness to previously described proteins as summarized in Table 1.

ORFX nucleic acids and polypeptides according to the invention can also be used to
10 identify cell types listed in Table 1 for an indicated ORFX according to the invention. Additional utilities for ORFX nucleic acids and polypeptides according to the invention are disclosed herein.

ORFX Nucleic Acids

15 The novel nucleic acids of the invention include those that encode an ORFX or ORFX-like protein, or biologically active portions thereof. The nucleic acids include nucleic acids encoding polypeptides that include the amino acid sequence of one or more of SEQ ID NO:2 n , wherein $n = 1$ to 3161. The encoded polypeptides can thus include, *e.g.*, the amino acid sequences of SEQ ID NO: 2, 4, 6, 8, 10, . . . , 6310, 6312, 6314, 6316, 6318, 6320, and/or 6322.

20 In some embodiments, a nucleic acid encoding a polypeptide having the amino acid sequence of one or more of SEQ ID NO:2 n (wherein $n = 1$ to 3161) includes the nucleic acid sequence of any of SEQ ID NO:2 $n-1$ (wherein $n = 1$ to 3161), or a fragment thereof. Additionally, the invention includes mutant or variant nucleic acids of any of SEQ ID NO:2 $n-1$ (wherein $n = 1$ to 3161), or a fragment thereof, any of whose bases may be changed from the
25 disclosed sequence while still encoding a protein that maintains its ORFX-like activities and physiological functions. The invention further includes the complement of the nucleic acid sequence of any of SEQ ID NO:2 $n-1$ (wherein $n = 1$ to 3161), including fragments, derivatives,

analog and homolog thereof. The invention additionally includes nucleic acids or nucleic acid fragments, or complements thereto, whose structures include chemical modifications.

Also included are nucleic acid fragments sufficient for use as hybridization probes to identify ORFX-encoding nucleic acids (*e.g.*, ORFX mRNA) and fragments for use as
5 polymerase chain reaction (PCR) primers for the amplification or mutation of ORFX nucleic acid molecules. As used herein, the term "nucleic acid molecule" is intended to include DNA molecules (*e.g.*, cDNA or genomic DNA), RNA molecules (*e.g.*, mRNA), analogs of the DNA or RNA generated using nucleotide analogs, and derivatives, fragments and homologs thereof. The nucleic acid molecule can be single-stranded or double-stranded, but preferably is
10 double-stranded DNA.

"Probes" refer to nucleic acid sequences of variable length, preferably between at least about 10 nucleotides (nt), 100 nt, or as many as about, *e.g.*, 6,000 nt, depending on use. Probes are used in the detection of identical, similar, or complementary nucleic acid sequences. Longer length probes are usually obtained from a natural or recombinant source, are highly specific and
15 much slower to hybridize than oligomers. Probes may be single- or double-stranded and designed to have specificity in PCR, membrane-based hybridization technologies, or ELISA-like technologies.

An "isolated" nucleic acid molecule is one that is separated from other nucleic acid molecules that are present in the natural source of the nucleic acid. Examples of isolated nucleic acid molecules include, but are not limited to, recombinant DNA molecules contained in a
20 vector, recombinant DNA molecules maintained in a heterologous host cell, partially or substantially purified nucleic acid molecules, and synthetic DNA or RNA molecules. Preferably, an "isolated" nucleic acid is free of sequences which naturally flank the nucleic acid (*i.e.*, sequences located at the 5' and 3' ends of the nucleic acid) in the genomic DNA of the organism
25 from which the nucleic acid is derived. For example, in various embodiments, the isolated ORFX nucleic acid molecule can contain less than about 50 kb, 25 kb, 5 kb, 4 kb, 3 kb, 2 kb, 1 kb, 0.5 kb or 0.1 kb of nucleotide sequences which naturally flank the nucleic acid molecule in genomic DNA of the cell from which the nucleic acid is derived. Moreover, an "isolated" nucleic acid molecule, such as a cDNA molecule, can be substantially free of other cellular
30 material or culture medium when produced by recombinant techniques, or of chemical precursors or other chemicals when chemically synthesized.

A nucleic acid molecule of the present invention, *e.g.*, a nucleic acid molecule having the nucleotide sequence of SEQ ID NO:2*n*-1 (wherein *n* = 1 to 3161), or a complement of any of this nucleotide sequence, can be isolated using standard molecular biology techniques and the sequence information provided herein. Using all or a portion of the nucleic acid sequence of any
5 of SEQ ID NO:2*n*-1 (wherein *n*=1 to 3161) as a hybridization probe, ORFX nucleic acid sequences can be isolated using standard hybridization and cloning techniques (*e.g.*, as described in Sambrook *et al.*, eds., MOLECULAR CLONING: A LABORATORY MANUAL 2nd Ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 1989; and Ausubel, *et al.*, eds., CURRENT PROTOCOLS IN MOLECULAR BIOLOGY, John Wiley & Sons, New York, NY, 1993.)

10 A nucleic acid of the invention can be amplified using cDNA, mRNA or alternatively, genomic DNA, as a template and appropriate oligonucleotide primers according to standard PCR amplification techniques. The nucleic acid so amplified can be cloned into an appropriate vector and characterized by DNA sequence analysis. Furthermore, oligonucleotides corresponding to ORFX nucleotide sequences can be prepared by standard synthetic techniques, *e.g.*, using an
15 automated DNA synthesizer.

As used herein, the term "oligonucleotide" refers to a series of linked nucleotide residues, which oligonucleotide has a sufficient number of nucleotide bases to be used in a PCR reaction. A short oligonucleotide sequence may be based on, or designed from, a genomic or cDNA
20 sequence and is used to amplify, confirm, or reveal the presence of an identical, similar or complementary DNA or RNA in a particular cell or tissue. Oligonucleotides comprise portions of a nucleic acid sequence having about 10 nt, 50 nt, or 100 nt in length, preferably about 15 nt to 30 nt in length. In one embodiment, an oligonucleotide comprising a nucleic acid molecule less than 100 nt in length would further comprise at least 6 contiguous nucleotides of any of
25 SEQ ID NO:2*n*-1 (wherein *n* = 1 to 3161), or a complement thereof. Oligonucleotides may be chemically synthesized and may be used as probes.

In another embodiment, an isolated nucleic acid molecule of the invention comprises a nucleic acid molecule that is a complement of the nucleotide sequence shown in any of SEQ ID NO:2*n*-1 (wherein *n* = 1 to 3161). In another embodiment, an isolated nucleic acid molecule of the invention comprises a nucleic acid molecule that is a complement of the nucleotide sequence
30 shown in any of SEQ ID NO:2*n*-1 (wherein *n* = 1 to 3161), or a portion of this nucleotide sequence. A nucleic acid molecule that is complementary to the nucleotide sequence shown in

is one that is sufficiently complementary to the nucleotide sequence shown in of any of SEQ ID NO:2*n*-1 (wherein *n* = 1 to 3161) that it can hydrogen bond with little or no mismatches to the nucleotide sequence shown in of any of SEQ ID NO:2*n*-1 (wherein *n* = 1 to 3161),, thereby forming a stable duplex.

- 5 As used herein, the term "complementary" refers to Watson-Crick or Hoogsteen base pairing between nucleotides units of a nucleic acid molecule, and the term "binding" means the physical or chemical interaction between two polypeptides or compounds or associated polypeptides or compounds or combinations thereof. Binding includes ionic, non-ionic, Von der Waals, hydrophobic interactions, etc. A physical interaction can be either direct or indirect.
- 10 Indirect interactions may be through or due to the effects of another polypeptide or compound. Direct binding refers to interactions that do not take place through, or due to, the effect of another polypeptide or compound, but instead are without other substantial chemical intermediates.

- Moreover, the nucleic acid molecule of the invention can comprise only a portion of the
- 15 nucleic acid sequence of any of SEQ ID NO:2*n*-1 (wherein *n* = 1 to 3161), *e.g.*, a fragment that can be used as a probe or primer, or a fragment encoding a biologically active portion of ORFX. Fragments provided herein are defined as sequences of at least 6 (contiguous) nucleic acids or at least 4 (contiguous) amino acids, a length sufficient to allow for specific hybridization in the case of nucleic acids or for specific recognition of an epitope in the case of amino acids, respectively,
- 20 and are at most some portion less than a full length sequence. Fragments may be derived from any contiguous portion of a nucleic acid or amino acid sequence of choice. Derivatives are nucleic acid sequences or amino acid sequences formed from the native compounds either directly or by modification or partial substitution. Analogs are nucleic acid sequences or amino acid sequences that have a structure similar to, but not identical to, the native compound but
- 25 differs from it in respect to certain components or side chains. Analogs may be synthetic or from a different evolutionary origin and may have a similar or opposite metabolic activity compared to wild type.

- Derivatives and analogs may be full length or other than full length, if the derivative or analog contains a modified nucleic acid or amino acid, as described below. Derivatives or
- 30 analogs of the nucleic acids or proteins of the invention include, but are not limited to, molecules comprising regions that are substantially homologous to the nucleic acids or proteins of the

invention, in various embodiments, by at least about 70%, 80%, 85%, 90%, 95%, 98%, or even 99% identity (with a preferred identity of 80-99%) over a nucleic acid or amino acid sequence of identical size or when compared to an aligned sequence in which the alignment is done by a computer homology program known in the art, or whose encoding nucleic acid is capable of hybridizing to the complement of a sequence encoding the aforementioned proteins under stringent, moderately stringent, or low stringent conditions. See *e.g.* Ausubel, *et al.*, CURRENT PROTOCOLS IN MOLECULAR BIOLOGY, John Wiley & Sons, New York, NY, 1993, and below. An exemplary program is the Gap program (Wisconsin Sequence Analysis Package, Version 8 for UNIX, Genetics Computer Group, University Research Park, Madison, WI) using the default settings, which uses the algorithm of Smith and Waterman (Adv. Appl. Math., 1981, 2: 482-489, which is incorporated herein by reference in its entirety).

A "homologous nucleic acid sequence" or "homologous amino acid sequence," or variations thereof, refer to sequences characterized by a homology at the nucleotide level or amino acid level as discussed above. Homologous nucleotide sequences encode those sequences coding for isoforms of ORFX polypeptide. Isoforms can be expressed in different tissues of the same organism as a result of, for example, alternative splicing of RNA. Alternatively, isoforms can be encoded by different genes. In the present invention, homologous nucleotide sequences include nucleotide sequences encoding for a ORFX polypeptide of species other than humans, including, but not limited to, mammals, and thus can include, *e.g.*, mouse, rat, rabbit, dog, cat, cow, horse, and other organisms. Homologous nucleotide sequences also include, but are not limited to, naturally occurring allelic variations and mutations of the nucleotide sequences set forth herein. A homologous nucleotide sequence does not, however, include the nucleotide sequence encoding human ORFX protein. Homologous nucleic acid sequences include those nucleic acid sequences that encode conservative amino acid substitutions (see below) in any of SEQ ID NO:2*n* (wherein *n* = 1 to 3161) as well as a polypeptide having ORFX activity. Biological activities of the ORFX proteins are described below. A homologous amino acid sequence does not encode the amino acid sequence of a human ORFX polypeptide.

The nucleotide sequence determined from the cloning of the human ORFX gene allows for the generation of probes and primers designed for use in identifying the cell types disclosed and/or cloning ORFX homologues in other cell types, *e.g.*, from other tissues, as well as ORFX homologues from other mammals. The probe/primer typically comprises a substantially purified

oligonucleotide. The oligonucleotide typically comprises a region of nucleotide sequence that hybridizes under stringent conditions to at least about 12, 25, 50, 100, 150, 200, 250, 300, 350 or 400 or more consecutive sense strand nucleotide sequence of SEQ ID NO:2*n*-1 (wherein *n* = 1 to 3161); or an anti-sense strand nucleotide sequence of SEQ ID NO:2*n*-1 (wherein *n* = 1 to 3161); or of a naturally occurring mutant of SEQ ID NO:2*n*-1 (wherein *n* = 1 to 3161).

Probes based on the human ORFX nucleotide sequence can be used to detect transcripts or genomic sequences encoding the same or homologous proteins. In various embodiments, the probe further comprises a label group attached thereto, *e.g.*, the label group can be a radioisotope, a fluorescent compound, an enzyme, or an enzyme co-factor. Such probes can be used as a part of a diagnostic test kit for identifying cells or tissue which misexpress a ORFX protein, such as by measuring a level of a ORFX-encoding nucleic acid in a sample of cells from a subject *e.g.*, detecting ORFX mRNA levels or determining whether a genomic ORFX gene has been mutated or deleted.

"A polypeptide having a biologically active portion of ORFX" refers to polypeptides exhibiting activity similar, but not necessarily identical to, an activity of a polypeptide of the present invention, including mature forms, as measured in a particular biological assay, with or without dose dependency. A nucleic acid fragment encoding a "biologically active portion of ORFX" can be prepared by isolating a portion of SEQ ID NO:2*n*-1 (wherein *n* = 1 to 3161), that encodes a polypeptide having a ORFX biological activity (biological activities of the ORFX proteins are summarized in Table 1), expressing the encoded portion of ORFX protein (*e.g.*, by recombinant expression *in vitro*) and assessing the activity of the encoded portion of ORFX. For example, a nucleic acid fragment encoding a biologically active portion of ORFX can optionally include a domain as shown in Table 1, column 4.

ORFX variants

The invention further encompasses nucleic acid molecules that differ from the disclosed ORFX nucleotide sequences due to degeneracy of the genetic code. These nucleic acids thus encode the same ORFX protein as that encoded by the nucleotide sequence shown in SEQ ID NO:2*n*-1 (wherein *n* = 1 to 3161). In another embodiment, an isolated nucleic acid molecule of the invention has a nucleotide sequence encoding a protein having an amino acid sequence shown in any of SEQ ID NO:2*n* (wherein *n* = 1 to 3161).

In addition to the human ORFX nucleotide sequence shown in any of SEQ ID NO:2*n*-1 (wherein *n* = 1 to 3161), it will be appreciated by those skilled in the art that DNA sequence polymorphisms that lead to changes in the amino acid sequences of ORFX may exist within a population (*e.g.*, the human population). Such genetic polymorphism in the ORFX gene may
5 exist among individuals within a population due to natural allelic variation. As used herein, the terms "gene" and "recombinant gene" refer to nucleic acid molecules comprising an open reading frame encoding a ORFX protein, preferably a mammalian ORFX protein. Such natural allelic variations can typically result in 1-5% variance in the nucleotide sequence of the ORFX gene. Any and all such nucleotide variations and resulting amino acid polymorphisms in ORFX that
10 are the result of natural allelic variation and that do not alter the functional activity of ORFX are intended to be within the scope of the invention.

Moreover, nucleic acid molecules encoding ORFX proteins from other species, and thus that have a nucleotide sequence that differs from the human sequence of any of SEQ ID NO:2*n*-1 (wherein *n* = 1 to 3161), are intended to be within the scope of the invention. Nucleic acid
15 molecules corresponding to natural allelic variants and homologues of the ORFX cDNAs of the invention can be isolated based on their homology to the human ORFX nucleic acids disclosed herein using the human cDNAs, or a portion thereof, as a hybridization probe according to standard hybridization techniques under stringent hybridization conditions.

In another embodiment, an isolated nucleic acid molecule of the invention is at least 6
20 nucleotides in length and hybridizes under stringent conditions to the nucleic acid molecule comprising the nucleotide sequence of any of SEQ ID NO:2*n*-1 (wherein *n* = 1 to 3161). In another embodiment, the nucleic acid is at least 10, 25, 50, 100, 250, 500 or 750 nucleotides in length. In another embodiment, an isolated nucleic acid molecule of the invention hybridizes to the coding region. As used herein, the term "hybridizes under stringent conditions" is intended
25 to describe conditions for hybridization and washing under which nucleotide sequences at least 60% homologous to each other typically remain hybridized to each other.

Homologs (*i.e.*, nucleic acids encoding ORFX proteins derived from species other than human) or other related sequences (*e.g.*, paralogs) can be obtained by low, moderate or high stringency hybridization with all or a portion of the particular human sequence as a probe using
30 methods well known in the art for nucleic acid hybridization and cloning.

As used herein, the phrase "stringent hybridization conditions" refers to conditions under which a probe, primer or oligonucleotide will hybridize to its target sequence, but to no other sequences. Stringent conditions are sequence-dependent and will be different in different circumstances. Longer sequences hybridize specifically at higher temperatures than shorter sequences. Generally, stringent conditions are selected to be about 5°C lower than the thermal melting point (T_m) for the specific sequence at a defined ionic strength and pH. The T_m is the temperature (under defined ionic strength, pH and nucleic acid concentration) at which 50% of the probes complementary to the target sequence hybridize to the target sequence at equilibrium. Since the target sequences are generally present at excess, at T_m , 50% of the probes are occupied at equilibrium. Typically, stringent conditions will be those in which the salt concentration is less than about 1.0 M sodium ion, typically about 0.01 to 1.0 M sodium ion (or other salts) at pH 7.0 to 8.3 and the temperature is at least about 30°C for short probes, primers or oligonucleotides (*e.g.*, 10 nt to 50 nt) and at least about 60°C for longer probes, primers and oligonucleotides. Stringent conditions may also be achieved with the addition of destabilizing agents, such as formamide.

Stringent conditions are known to those skilled in the art and can be found in CURRENT PROTOCOLS IN MOLECULAR BIOLOGY, John Wiley & Sons, N.Y. (1989), 6.3.1-6.3.6. Preferably, the conditions are such that sequences at least about 65%, 70%, 75%, 85%, 90%, 95%, 98%, or 99% homologous to each other typically remain hybridized to each other. A non-limiting example of stringent hybridization conditions is hybridization in a high salt buffer comprising 6X SSC, 50 mM Tris-HCl (pH 7.5), 1 mM EDTA, 0.02% PVP, 0.02% Ficoll, 0.02% BSA, and 500 mg/ml denatured salmon sperm DNA at 65°C. This hybridization is followed by one or more washes in 0.2X SSC, 0.01% BSA at 50°C. An isolated nucleic acid molecule of the invention that hybridizes under stringent conditions to the sequence of any of SEQ ID NO:2*n*-1 (wherein $n = 1$ to 3161) corresponds to a naturally occurring nucleic acid molecule. As used herein, a "naturally-occurring" nucleic acid molecule refers to an RNA or DNA molecule having a nucleotide sequence that occurs in nature (*e.g.*, encodes a natural protein).

In a second embodiment, a nucleic acid sequence that is hybridizable to the nucleic acid molecule comprising the nucleotide sequence of any of SEQ ID NO:2*n*-1 (wherein $n = 1$ to 3161), or fragments, analogs or derivatives thereof, under conditions of moderate stringency is provided. A non-limiting example of moderate stringency hybridization conditions are

hybridization in 6X SSC, 5X Denhardt's solution, 0.5% SDS and 100 mg/ml denatured salmon sperm DNA at 55°C, followed by one or more washes in 1X SSC, 0.1% SDS at 37°C. Other conditions of moderate stringency that may be used are well known in the art. See, *e.g.*, Ausubel *et al.* (eds.), 1993, CURRENT PROTOCOLS IN MOLECULAR BIOLOGY, John Wiley & Sons, NY, and
5 Kriegler, 1990, GENE TRANSFER AND EXPRESSION, A LABORATORY MANUAL, Stockton Press, NY.

In a third embodiment, a nucleic acid that is hybridizable to the nucleic acid molecule comprising the nucleotide sequence of any of SEQ ID NO:2*n*-1 (wherein *n* = 1 to 3161), or fragments, analogs or derivatives thereof, under conditions of low stringency, is provided. A
10 non-limiting example of low stringency hybridization conditions are hybridization in 35% formamide, 5X SSC, 50 mM Tris-HCl (pH 7.5), 5 mM EDTA, 0.02% PVP, 0.02% Ficoll, 0.2% BSA, 100 mg/ml denatured salmon sperm DNA, 10% (wt/vol) dextran sulfate at 40°C, followed by one or more washes in 2X SSC, 25 mM Tris-HCl (pH 7.4), 5 mM EDTA, and 0.1% SDS at
-- 50°C. Other conditions of low stringency that may be used are well known in the art (*e.g.*, as
15 employed for cross-species hybridizations). See, *e.g.*, Ausubel *et al.* (eds.), 1993, CURRENT PROTOCOLS IN MOLECULAR BIOLOGY, John Wiley & Sons, NY, and Kriegler, 1990, GENE TRANSFER AND EXPRESSION, A LABORATORY MANUAL, Stockton Press, NY; Shilo and Weinberg, 1981, *Proc Natl Acad Sci USA* 78: 6789-6792.

Conservative mutations

20 In addition to naturally-occurring allelic variants of the ORFX sequence that may exist in the population, the skilled artisan will further appreciate that changes can be introduced by mutation into the nucleotide sequence of any of SEQ ID NO:2*n*-1 (wherein *n* = 1 to 3161), thereby leading to changes in the amino acid sequence of the encoded ORFX protein, without altering the functional ability of the ORFX protein. For example, nucleotide substitutions
25 leading to amino acid substitutions at "non-essential" amino acid residues can be made in the sequence of any of SEQ ID NO:2*n*-1 (wherein *n* = 1 to 3161). A "non-essential" amino acid residue is a residue that can be altered from the wild-type sequence of ORFX without altering the biological activity, whereas an "essential" amino acid residue is required for biological activity. For example, amino acid residues that are conserved among the ORFX proteins of the present
30 invention, are predicted to be particularly unamenable to alteration.

Amino acid residues that are conserved among members of an ORFX family members are predicted to be less amenable to alteration. For example, an ORFX protein according to the present invention can contain at least one domain (*e.g.*, as shown in Table 1) that is a typically conserved region in an ORFX family member. As such, these conserved domains are not likely to be amenable to mutation. Other amino acid residues, however, (*e.g.*, those that are not conserved or only semi-conserved among members of the ORFX family) may not be as essential for activity and thus are more likely to be amenable to alteration.

Another aspect of the invention pertains to nucleic acid molecules encoding ORFX proteins that contain changes in amino acid residues that are not essential for activity. Such ORFX proteins differ in amino acid sequence from any of any of SEQ ID NO:2*n* (wherein *n* = 1 to 3161), yet retain biological activity. In one embodiment, the isolated nucleic acid molecule comprises a nucleotide sequence encoding a protein, wherein the protein comprises an amino acid sequence at least about 75% homologous to the amino acid sequence of any of SEQ ID NO:2*n* (wherein *n* = 1 to 3161). Preferably, the protein encoded by the nucleic acid is at least about 80% homologous to any of SEQ ID NO:2*n* (wherein *n* = 1 to 3161), more preferably at least about 90%, 95%, 98%, and most preferably at least about 99% homologous to SEQ ID NO:2.

An isolated nucleic acid molecule encoding a ORFX protein homologous to the protein of any of SEQ ID NO:2*n* (wherein *n* = 1 to 3161) can be created by introducing one or more nucleotide substitutions, additions or deletions into the corresponding nucleotide sequence, *i.e.* SEQ ID NO:2*n*-1 for the corresponding *n*, such that one or more amino acid substitutions, additions or deletions are introduced into the encoded protein.

Mutations can be introduced into SEQ ID NO:2*n*-1 (wherein *n* = 1 to 3161) by standard techniques, such as site-directed mutagenesis and PCR-mediated mutagenesis. Preferably, conservative amino acid substitutions are made at one or more predicted non-essential amino acid residues. A "conservative amino acid substitution" is one in which the amino acid residue is replaced with an amino acid residue having a similar side chain. Families of amino acid residues having similar side chains have been defined in the art. These families include amino acids with basic side chains (*e.g.*, lysine, arginine, histidine), acidic side chains (*e.g.*, aspartic acid, glutamic acid), uncharged polar side chains (*e.g.*, glycine, asparagine, glutamine, serine, threonine, tyrosine, cysteine), nonpolar side chains (*e.g.*, alanine, valine, leucine, isoleucine, proline,

phenylalanine, methionine, tryptophan), beta-branched side chains (*e.g.*, threonine, valine, isoleucine) and aromatic side chains (*e.g.*, tyrosine, phenylalanine, tryptophan, histidine). Thus, a predicted nonessential amino acid residue in ORFX is replaced with another amino acid residue from the same side chain family. Alternatively, in another embodiment, mutations can be
5 introduced randomly along all or part of a ORFX coding sequence, such as by saturation mutagenesis, and the resultant mutants can be screened for ORFX biological activity to identify mutants that retain activity. Following mutagenesis of SEQ ID NO:2*n*-1 (wherein *n* = 1 to 3161), the encoded protein can be expressed by any recombinant technology known in the art and the activity of the protein can be determined.

10 In one embodiment, a mutant ORFX protein can be assayed for (1) the ability to form protein:protein interactions with other ORFX proteins, other cell-surface proteins, or biologically active portions thereof, (2) complex formation between a mutant ORFX protein and a ORFX receptor; (3) the ability of a mutant ORFX protein to bind to an intracellular target protein or biologically active portion thereof; (*e.g.*, avidin proteins); (4) the ability to bind BRA protein; or
15 (5) the ability to specifically bind an anti-ORFX protein antibody.

Antisense

Another aspect of the invention pertains to isolated antisense nucleic acid molecules that are hybridizable to or complementary to the nucleic acid molecule comprising the nucleotide sequence of SEQ ID NO:2*n*-1 (wherein *n* = 1 to 3161), or fragments, analogs or derivatives
20 thereof. An "antisense" nucleic acid comprises a nucleotide sequence that is complementary to a "sense" nucleic acid encoding a protein, *e.g.*, complementary to the coding strand of a double-stranded cDNA molecule or complementary to an mRNA sequence. In specific aspects, antisense nucleic acid molecules are provided that comprise a sequence complementary to at least about 10, 25, 50, 100, 250 or 500 nucleotides or an entire ORFX coding strand, or to only a
25 portion thereof. Nucleic acid molecules encoding fragments, homologs, derivatives and analogs of a ORFX protein of any of SEQ ID NO:2*n* (wherein *n* = 1 to 3161) or antisense nucleic acids complementary to a ORFX nucleic acid sequence of SEQ ID NO:2*n*-1 (wherein *n* = 1 to 3161) are additionally provided.

30 In one embodiment, an antisense nucleic acid molecule is antisense to a "coding region" of the coding strand of a nucleotide sequence encoding ORFX. The term "coding region" refers to the region of the nucleotide sequence comprising codons which are translated into amino acid

residues (*e.g.*, the protein coding region of a human ORFX that corresponds to any of SEQ ID NO:2*n* (wherein *n* = 1 to 3161)). In another embodiment, the antisense nucleic acid molecule is antisense to a "noncoding region" of the coding strand of a nucleotide sequence encoding ORFX. The term "noncoding region" refers to 5' and 3' sequences which flank the coding region that are not translated into amino acids (*i.e.*, also referred to as 5' and 3' untranslated regions).

Given the coding strand sequences encoding ORFX disclosed herein (*e.g.*, SEQ ID NO:2*n*-1 (wherein *n* = 1 to 3161)), antisense nucleic acids of the invention can be designed according to the rules of Watson and Crick or Hoogsteen base pairing. The antisense nucleic acid molecule can be complementary to the entire coding region of ORFX mRNA, but more preferably is an oligonucleotide that is antisense to only a portion of the coding or noncoding region of ORFX mRNA. For example, the antisense oligonucleotide can be complementary to the region surrounding the translation start site of ORFX mRNA. An antisense oligonucleotide can be, for example, about 5, 10, 15, 20, 25, 30, 35, 40, 45 or 50 nucleotides in length. An antisense nucleic acid of the invention can be constructed using chemical synthesis or enzymatic ligation reactions using procedures known in the art. For example, an antisense nucleic acid (*e.g.*, an antisense oligonucleotide) can be chemically synthesized using naturally occurring nucleotides or variously modified nucleotides designed to increase the biological stability of the molecules or to increase the physical stability of the duplex formed between the antisense and sense nucleic acids, *e.g.*, phosphorothioate derivatives and acridine substituted nucleotides can be used.

Examples of modified nucleotides that can be used to generate the antisense nucleic acid include: 5-fluorouracil, 5-bromouracil, 5-chlorouracil, 5-iodouracil, hypoxanthine, xanthine, 4-acetylcytosine, 5-(carboxyhydroxymethyl) uracil, 5-carboxymethylaminomethyl-2-thiouridine, 5-carboxymethylaminomethyluracil, dihydrouracil, beta-D-galactosylqueosine, inosine, N6-isopentenyladenine, 1-methylguanine, 1-methylinosine, 2,2-dimethylguanine, 2-methyladenine, 2-methylguanine, 3-methylcytosine, 5-methylcytosine, N6-adenine, 7-methylguanine, 5-methylaminomethyluracil, 5-methoxyaminomethyl-2-thiouracil, beta-D-mannosylqueosine, 5'-methoxycarboxymethyluracil, 5-methoxyuracil, 2-methylthio-N6-isopentenyladenine, uracil-5-oxyacetic acid (*v*), wybutoxosine, pseudouracil, queosine, 2-thiocytosine, 5-methyl-2-thiouracil, 2-thiouracil, 4-thiouracil, 5-methyluracil, uracil-5-oxyacetic acid methylester, uracil-5-oxyacetic acid (*v*), 5-methyl-2-thiouracil,

3-(3-amino-3-N-2-carboxypropyl) uracil, (acp3)w, and 2,6-diaminopurine. Alternatively, the antisense nucleic acid can be produced biologically using an expression vector into which a nucleic acid has been subcloned in an antisense orientation (*i.e.*, RNA transcribed from the inserted nucleic acid will be of an antisense orientation to a target nucleic acid of interest, described further in the following subsection).

The antisense nucleic acid molecules of the invention are typically administered to a subject or generated *in situ* such that they hybridize with or bind to cellular mRNA and/or genomic DNA encoding a ORFX protein to thereby inhibit expression of the protein, *e.g.*, by inhibiting transcription and/or translation. The hybridization can be by conventional nucleotide complementarity to form a stable duplex, or, for example, in the case of an antisense nucleic acid molecule that binds to DNA duplexes, through specific interactions in the major groove of the double helix. An example of a route of administration of antisense nucleic acid molecules of the invention includes direct injection at a tissue site. Alternatively, antisense nucleic acid molecules can be modified to target selected cells and then administered systemically. For example, for systemic administration, antisense molecules can be modified such that they specifically bind to receptors or antigens expressed on a selected cell surface, *e.g.*, by linking the antisense nucleic acid molecules to peptides or antibodies that bind to cell surface receptors or antigens. The antisense nucleic acid molecules can also be delivered to cells using the vectors described herein. To achieve sufficient intracellular concentrations of antisense molecules, vector constructs in which the antisense nucleic acid molecule is placed under the control of a strong pol II or pol III promoter are preferred.

In yet another embodiment, the antisense nucleic acid molecule of the invention is an α -anomeric nucleic acid molecule. An α -anomeric nucleic acid molecule forms specific double-stranded hybrids with complementary RNA in which, contrary to the usual β -units, the strands run parallel to each other (Gaultier *et al.* (1987) *Nucleic Acids Res* 15: 6625-6641). The antisense nucleic acid molecule can also comprise a 2'-o-methylribonucleotide (Inoue *et al.* (1987) *Nucleic Acids Res* 15: 6131-6148) or a chimeric RNA -DNA analogue (Inoue *et al.* (1987) *FEBS Lett* 215: 327-330).

Ribozymes and PNA moieties

Such modifications include, by way of nonlimiting example, modified bases, and nucleic acids whose sugar phosphate backbones are modified or derivatized. These modifications are

carried out at least in part to enhance the chemical stability of the modified nucleic acid, such that they may be used, for example, as antisense binding nucleic acids in therapeutic applications in a subject.

In still another embodiment, an antisense nucleic acid of the invention is a ribozyme.

- 5 Ribozymes are catalytic RNA molecules with ribonuclease activity that are capable of cleaving a single-stranded nucleic acid, such as an mRNA, to which they have a complementary region. Thus, ribozymes (*e.g.*, hammerhead ribozymes (described in Haselhoff and Gerlach (1988) *Nature* 334:585-591)) can be used to catalytically cleave ORFX mRNA transcripts to thereby inhibit translation of ORFX mRNA. A ribozyme having specificity for a ORFX-encoding
- 10 nucleic acid can be designed based upon the nucleotide sequence of a ORFX DNA disclosed herein (*i.e.*, SEQ ID NO:2*n*-1 (wherein *n* = 1 to 3161)). For example, a derivative of a Tetrahymena L-19 IVS RNA can be constructed in which the nucleotide sequence of the active site is complementary to the nucleotide sequence to be cleaved in a ORFX-encoding mRNA. See, *e.g.*, Cech *et al.* U.S. Pat. No. 4,987,071; and Cech *et al.* U.S. Pat. No. 5,116,742.
- 15 Alternatively, ORFX mRNA can be used to select a catalytic RNA having a specific ribonuclease activity from a pool of RNA molecules. See, *e.g.*, Bartel *et al.*, (1993) *Science* 261:1411-1418.

- Alternatively, ORFX gene expression can be inhibited by targeting nucleotide sequences complementary to the regulatory region of the ORFX (*e.g.*, the ORFX promoter and/or
- 20 enhancers) to form triple helical structures that prevent transcription of the ORFX gene in target cells. See generally, Helene. (1991) *Anticancer Drug Des.* 6: 569-84; Helene. *et al.* (1992) *Ann. N.Y. Acad. Sci.* 660:27-36; and Maher (1992) *Bioassays* 14: 807-15.

- In various embodiments, the nucleic acids of ORFX can be modified at the base moiety, sugar moiety or phosphate backbone to improve, *e.g.*, the stability, hybridization, or solubility of
- 25 the molecule. For example, the deoxyribose phosphate backbone of the nucleic acids can be modified to generate peptide nucleic acids (see Hyrup *et al.* (1996) *Bioorg Med Chem* 4: 5-23). As used herein, the terms "peptide nucleic acids" or "PNAs" refer to nucleic acid mimics, *e.g.*, DNA mimics, in which the deoxyribose phosphate backbone is replaced by a pseudopeptide backbone and only the four natural nucleobases are retained. The neutral backbone of PNAs has
- 30 been shown to allow for specific hybridization to DNA and RNA under conditions of low ionic strength. The synthesis of PNA oligomers can be performed using standard solid phase peptide

synthesis protocols as described in Hyrup *et al.* (1996) above; Perry-O'Keefe *et al.* (1996) *PNAS* 93: 14670-675.

PNAs of ORFX can be used in therapeutic and diagnostic applications. For example, PNAs can be used as antisense or antigene agents for sequence-specific modulation of gene expression by, *e.g.*, inducing transcription or translation arrest or inhibiting replication. PNAs of ORFX can also be used, *e.g.*, in the analysis of single base pair mutations in a gene by, *e.g.*, PNA directed PCR clamping; as artificial restriction enzymes when used in combination with other enzymes, *e.g.*, S1 nucleases (Hyrup B. (1996) above); or as probes or primers for DNA sequence and hybridization (Hyrup *et al.* (1996), above; Perry-O'Keefe (1996), above).

In another embodiment, PNAs of ORFX can be modified, *e.g.*, to enhance their stability or cellular uptake, by attaching lipophilic or other helper groups to PNA, by the formation of PNA-DNA chimeras, or by the use of liposomes or other techniques of drug delivery known in the art. For example, PNA-DNA chimeras of ORFX can be generated that may combine the advantageous properties of PNA and DNA. Such chimeras allow DNA recognition enzymes, *e.g.*, RNase H and DNA polymerases, to interact with the DNA portion while the PNA portion would provide high binding affinity and specificity. PNA-DNA chimeras can be linked using linkers of appropriate lengths selected in terms of base stacking, number of bonds between the nucleobases, and orientation (Hyrup (1996) above). The synthesis of PNA-DNA chimeras can be performed as described in Hyrup (1996) above and Finn *et al.* (1996) *Nucl Acids Res* 24: 3357-63. For example, a DNA chain can be synthesized on a solid support using standard phosphoramidite coupling chemistry, and modified nucleoside analogs, *e.g.*, 5'-(4-methoxytrityl)amino-5'-deoxy-thymidine phosphoramidite, can be used between the PNA and the 5' end of DNA (Mag *et al.* (1989) *Nucl Acid Res* 17: 5973-88). PNA monomers are then coupled in a stepwise manner to produce a chimeric molecule with a 5' PNA segment and a 3' DNA segment (Finn *et al.* (1996) above). Alternatively, chimeric molecules can be synthesized with a 5' DNA segment and a 3' PNA segment. See, Petersen *et al.* (1975) *Bioorg Med Chem Lett* 5: 1119-11124.

In other embodiments, the oligonucleotide may include other appended groups such as peptides (*e.g.*, for targeting host cell receptors *in vivo*), or agents facilitating transport across the cell membrane (see, *e.g.*, Letsinger *et al.*, 1989, *Proc. Natl. Acad. Sci. U.S.A.* 86:6553-6556; Lemaitre *et al.*, 1987, *Proc. Natl. Acad. Sci.* 84:648-652; PCT Publication No. W088/09810) or

the blood-brain barrier (see, *e.g.*, PCT Publication No. W089/10134). In addition, oligonucleotides can be modified with hybridization triggered cleavage agents (See, *e.g.*, Krol *et al.*, 1988, *BioTechniques* 6:958-976) or intercalating agents. (See, *e.g.*, Zon, 1988, *Pharm. Res.* 5: 539-549). To this end, the oligonucleotide may be conjugated to another molecule, *e.g.*, a peptide, a hybridization triggered cross-linking agent, a transport agent, a hybridization-triggered cleavage agent, etc.

ORFX polypeptides

The novel protein of the invention includes the ORFX-like protein whose sequence is provided in any of SEQ ID NO:2*n* (wherein *n* = 1 to 3161). The invention also includes a mutant or variant protein any of whose residues may be changed from the corresponding residue shown in FIG. 1 while still encoding a protein that maintains its ORFX-like activities and physiological functions, or a functional fragment thereof. For example, the invention includes the polypeptides encoded by the variant ORFX nucleic acids described above. In the mutant or variant protein, up to 20% or more of the residues may be so changed.

In general, an ORFX-like variant that preserves ORFX-like function includes any variant in which residues at a particular position in the sequence have been substituted by other amino acids, and further include the possibility of inserting an additional residue or residues between two residues of the parent protein as well as the possibility of deleting one or more residues from the parent sequence. Any amino acid substitution, insertion, or deletion is encompassed by the invention. In favorable circumstances, the substitution is a conservative substitution as defined above. Furthermore, without limiting the scope of the invention, positions of any of SEQ ID NO:2*n* (wherein *n* = 1 to 3161) may be substituted such that a mutant or variant protein may include one or more substitutions

The invention also includes isolated ORFX proteins, and biologically active portions thereof, or derivatives, fragments, analogs or homologs thereof. Also provided are polypeptide fragments suitable for use as immunogens to raise anti-ORFX antibodies. In one embodiment, native ORFX proteins can be isolated from cells or tissue sources by an appropriate purification scheme using standard protein purification techniques. In another embodiment, ORFX proteins are produced by recombinant DNA techniques. Alternative to recombinant expression, a ORFX

protein or polypeptide can be synthesized chemically using standard peptide synthesis techniques.

An "isolated" or "purified" protein or biologically active portion thereof is substantially free of cellular material or other contaminating proteins from the cell or tissue source from which the ORFX protein is derived, or substantially free from chemical precursors or other chemicals when chemically synthesized. The language "substantially free of cellular material" includes preparations of ORFX protein in which the protein is separated from cellular components of the cells from which it is isolated or recombinantly produced. In one embodiment, the language "substantially free of cellular material" includes preparations of ORFX protein having less than about 30% (by dry weight) of non-ORFX protein (also referred to herein as a "contaminating protein"), more preferably less than about 20% of non-ORFX protein, still more preferably less than about 10% of non-ORFX protein, and most preferably less than about 5% non-ORFX protein. When the ORFX protein or biologically active portion thereof is recombinantly produced, it is also preferably substantially free of culture medium, *i.e.*, culture medium represents less than about 20%, more preferably less than about 10%, and most preferably less than about 5% of the volume of the protein preparation.

The language "substantially free of chemical precursors or other chemicals" includes preparations of ORFX protein in which the protein is separated from chemical precursors or other chemicals that are involved in the synthesis of the protein. In one embodiment, the language "substantially free of chemical precursors or other chemicals" includes preparations of ORFX protein having less than about 30% (by dry weight) of chemical precursors or non-ORFX chemicals, more preferably less than about 20% chemical precursors or non-ORFX chemicals, still more preferably less than about 10% chemical precursors or non-ORFX chemicals, and most preferably less than about 5% chemical precursors or non-ORFX chemicals.

Biologically active portions of a ORFX protein include peptides comprising amino acid sequences sufficiently homologous to or derived from the amino acid sequence of the ORFX protein, *e.g.*, the amino acid sequence shown in SEQ ID NO:2 that include fewer amino acids than the full length ORFX proteins, and exhibit at least one activity of a ORFX protein. Typically, biologically active portions comprise a domain or motif with at least one activity of the ORFX protein. A biologically active portion of a ORFX protein can be a polypeptide which is, for example, 10, 25, 50, 100 or more amino acids in length.

A biologically active portion of a ORFX protein of the present invention may contain at least one of the above-identified domains conserved between the FGF family of proteins. Moreover, other biologically active portions, in which other regions of the protein are deleted, can be prepared by recombinant techniques and evaluated for one or more of the functional activities of a native ORFX protein.

In an embodiment, the ORFX protein has an amino acid sequence shown in any of SEQ ID NO:2n (wherein n = 1 to 3161). In other embodiments, the ORFX protein is substantially homologous to any of SEQ ID NO:2n (wherein n = 1 to 3161) and retains the functional activity of the protein of any of SEQ ID NO:2n (wherein n = 1 to 3161), yet differs in amino acid sequence due to natural allelic variation or mutagenesis, as described in detail below.

Accordingly, in another embodiment, the ORFX protein is a protein that comprises an amino acid sequence at least about 45% homologous, and more preferably about 55, 65, 70, 75, 80, 85, 90, 95, 98 or even 99% homologous to the amino acid sequence of any of SEQ ID NO:2n (wherein n = 1 to 3161) and retains the functional activity of the ORFX proteins of the corresponding polypeptide having the sequence of SEQ ID NO:2n (wherein n = 1 to 3161).

Determining homology between two or more sequences

To determine the percent homology of two amino acid sequences or of two nucleic acids, the sequences are aligned for optimal comparison purposes (*e.g.*, gaps can be introduced in either of the sequences being compared for optimal alignment between the sequences). The amino acid residues or nucleotides at corresponding amino acid positions or nucleotide positions are then compared. When a position in the first sequence is occupied by the same amino acid residue or nucleotide as the corresponding position in the second sequence, then the molecules are homologous at that position (*i.e.*, as used herein amino acid or nucleic acid "homology" is equivalent to amino acid or nucleic acid "identity").

The nucleic acid sequence homology may be determined as the degree of identity between two sequences. The homology may be determined using computer programs known in the art, such as GAP software provided in the GCG program package. See, *Needleman and Wunsch* 1970 *J Mol Biol* 48: 443-453. Using GCG GAP software with the following settings for nucleic acid sequence comparison: GAP creation penalty of 5.0 and GAP extension penalty of 0.3, the coding region of the analogous nucleic acid sequences referred to above exhibits a

degree of identity preferably of at least 70%, 75%, 80%, 85%, 90%, 95%, 98%, or 99%, with the CDS (encoding) part of the DNA sequence shown in SEQ ID NO:2n-1 (wherein n = 1 to 3161).

The term "sequence identity" refers to the degree to which two polynucleotide or polypeptide sequences are identical on a residue-by-residue basis over a particular region of comparison. The term "percentage of sequence identity" is calculated by comparing two optimally aligned sequences over that region of comparison, determining the number of positions at which the identical nucleic acid base (*e.g.*, A, T, C, G, U, or I, in the case of nucleic acids) occurs in both sequences to yield the number of matched positions, dividing the number of matched positions by the total number of positions in the region of comparison (*i.e.*, the window size), and multiplying the result by 100 to yield the percentage of sequence identity. The term "substantial identity" as used herein denotes a characteristic of a polynucleotide sequence, wherein the polynucleotide comprises a sequence that has at least 80 percent sequence identity, preferably at least 85 percent identity and often 90 to 95 percent sequence identity, more usually at least 99 percent sequence identity as compared to a reference sequence over a comparison region. The term "percentage of positive residues" is calculated by comparing two optimally aligned sequences over that region of comparison, determining the number of positions at which the identical and conservative amino acid substitutions, as defined above, occur in both sequences to yield the number of matched positions, dividing the number of matched positions by the total number of positions in the region of comparison (*i.e.*, the window size), and multiplying the result by 100 to yield the percentage of positive residues.

Chimeric and fusion proteins

The invention also provides ORFX chimeric or fusion proteins. As used herein, a ORFX "chimeric protein" or "fusion protein" includes a ORFX polypeptide operatively linked to a non-ORFX polypeptide. A "ORFX polypeptide" refers to a polypeptide having an amino acid sequence corresponding to ORFX, whereas a "non-ORFX polypeptide" refers to a polypeptide having an amino acid sequence corresponding to a protein that is not substantially homologous to the ORFX protein, *e.g.*, a protein that is different from the ORFX protein and that is derived from the same or a different organism. Within a ORFX fusion protein the ORFX polypeptide can correspond to all or a portion of a ORFX protein. In one embodiment, a ORFX fusion protein comprises at least one biologically active portion of a ORFX protein. In another embodiment, a ORFX fusion protein comprises at least two biologically active portions of a

ORFX protein. Within the fusion protein, the term "operatively linked" is intended to indicate that the ORFX polypeptide and the non-ORFX polypeptide are fused in-frame to each other. The non-ORFX polypeptide can be fused to the N-terminus or C-terminus of the ORFX polypeptide.

5 For example, in one embodiment a ORFX fusion protein comprises a ORFX polypeptide operably linked to the extracellular domain of a second protein. Such fusion proteins can be further utilized in screening assays for compounds that modulate ORFX activity (such assays are described in detail below).

10 In another embodiment, the fusion protein is a GST-ORFX fusion protein in which the ORFX sequences are fused to the C-terminus of the GST (i.e., glutathione S-transferase) sequences. Such fusion proteins can facilitate the purification of recombinant ORFX.

15 In yet another embodiment, the fusion protein is a ORFX protein containing a heterologous signal sequence at its N-terminus. For example, the native ORFX signal sequence can be removed and replaced with a signal sequence from another protein. In certain host cells (e.g., mammalian host cells), expression and/or secretion of ORFX can be increased through use of a heterologous signal sequence.

20 In another embodiment, the fusion protein is a ORFX-immunoglobulin fusion protein in which the ORFX sequences comprising one or more domains are fused to sequences derived from a member of the immunoglobulin protein family. The ORFX-immunoglobulin fusion proteins of the invention can be incorporated into pharmaceutical compositions and administered to a subject to inhibit an interaction between a ORFX ligand and a ORFX protein on the surface of a cell, to thereby suppress ORFX-mediated signal transduction *in vivo*. In one nonlimiting example, a contemplated ORFX ligand of the invention is an ORFX receptor. The ORFX-immunoglobulin fusion proteins can be used to modulate the bioavailability of a ORFX cognate ligand. Inhibition of the ORFX ligand/ORFX interaction may be useful therapeutically for both the treatment of proliferative and differentiative disorders, as well as modulating (e.g., promoting or inhibiting) cell survival. Moreover, the ORFX-immunoglobulin fusion proteins of the invention can be used as immunogens to produce anti-ORFX antibodies in a subject, to purify ORFX ligands, and in screening assays to identify molecules that inhibit the interaction of
30 ORFX with a ORFX ligand.

A ORFX chimeric or fusion protein of the invention can be produced by standard recombinant DNA techniques. For example, DNA fragments coding for the different polypeptide sequences are ligated together in-frame in accordance with conventional techniques, e.g., by employing blunt-ended or stagger-ended termini for ligation, restriction enzyme digestion to provide for appropriate termini, filling-in of cohesive ends as appropriate, alkaline phosphatase treatment to avoid undesirable joining, and enzymatic ligation. In another embodiment, the fusion gene can be synthesized by conventional techniques including automated DNA synthesizers. Alternatively, PCR amplification of gene fragments can be carried out using anchor primers that give rise to complementary overhangs between two consecutive gene fragments that can subsequently be annealed and reamplified to generate a chimeric gene sequence (see, for example, Ausubel et al. (eds.) CURRENT PROTOCOLS IN MOLECULAR BIOLOGY, John Wiley & Sons, 1992). Moreover, many expression vectors are commercially available that already encode a fusion moiety (e.g., a GST polypeptide). A ORFX-encoding nucleic acid can be cloned into such an expression vector such that the fusion moiety is linked in-frame to the ORFX protein.

ORFX agonists and antagonists

The present invention also pertains to variants of the ORFX proteins that function as either ORFX agonists (mimetics) or as ORFX antagonists. Variants of the ORFX protein can be generated by mutagenesis, e.g., discrete point mutation or truncation of the ORFX protein. An agonist of the ORFX protein can retain substantially the same, or a subset of, the biological activities of the naturally occurring form of the ORFX protein. An antagonist of the ORFX protein can inhibit one or more of the activities of the naturally occurring form of the ORFX protein by, for example, competitively binding to a downstream or upstream member of a cellular signaling cascade which includes the ORFX protein. Thus, specific biological effects can be elicited by treatment with a variant of limited function. In one embodiment, treatment of a subject with a variant having a subset of the biological activities of the naturally occurring form of the protein has fewer side effects in a subject relative to treatment with the naturally occurring form of the ORFX proteins.

Variants of the ORFX protein that function as either ORFX agonists (mimetics) or as ORFX antagonists can be identified by screening combinatorial libraries of mutants, e.g., truncation mutants, of the ORFX protein for ORFX protein agonist or antagonist activity. In one

embodiment, a variegated library of ORFX variants is generated by combinatorial mutagenesis at the nucleic acid level and is encoded by a variegated gene library. A variegated library of ORFX variants can be produced by, for example, enzymatically ligating a mixture of synthetic oligonucleotides into gene sequences such that a degenerate set of potential ORFX sequences is expressible as individual polypeptides, or alternatively, as a set of larger fusion proteins (e.g., for phage display) containing the set of ORFX sequences therein. There are a variety of methods which can be used to produce libraries of potential ORFX variants from a degenerate oligonucleotide sequence. Chemical synthesis of a degenerate gene sequence can be performed in an automatic DNA synthesizer, and the synthetic gene then ligated into an appropriate expression vector. Use of a degenerate set of genes allows for the provision, in one mixture, of all of the sequences encoding the desired set of potential ORFX sequences. Methods for synthesizing degenerate oligonucleotides are known in the art (see, e.g., Narang (1983) *Tetrahedron* 39:3; Itakura *et al.* (1984) *Annu Rev Biochem* 53:323; Itakura *et al.* (1984) *Science* 198:1056; Ike *et al.* (1983) *Nucl Acid Res* 11:477.

15 Polypeptide libraries

In addition, libraries of fragments of the ORFX protein coding sequence can be used to generate a variegated population of ORFX fragments for screening and subsequent selection of variants of a ORFX protein. In one embodiment, a library of coding sequence fragments can be generated by treating a double stranded PCR fragment of a ORFX coding sequence with a nuclease under conditions wherein nicking occurs only about once per molecule, denaturing the double stranded DNA, renaturing the DNA to form double stranded DNA that can include sense/antisense pairs from different nicked products, removing single stranded portions from reformed duplexes by treatment with S1 nuclease, and ligating the resulting fragment library into an expression vector. By this method, an expression library can be derived which encodes N-terminal and internal fragments of various sizes of the ORFX protein.

Several techniques are known in the art for screening gene products of combinatorial libraries made by point mutations or truncation, and for screening cDNA libraries for gene products having a selected property. Such techniques are adaptable for rapid screening of the gene libraries generated by the combinatorial mutagenesis of ORFX proteins. The most widely used techniques, which are amenable to high throughput analysis, for screening large gene libraries typically include cloning the gene library into replicable expression vectors,

transforming appropriate cells with the resulting library of vectors, and expressing the combinatorial genes under conditions in which detection of a desired activity facilitates isolation of the vector encoding the gene whose product was detected. Recursive ensemble mutagenesis (REM), a new technique that enhances the frequency of functional mutants in the libraries, can
5 be used in combination with the screening assays to identify ORFX variants (Arkin and Yourvan (1992) PNAS 89:7811-7815; Delgrave *et al.* (1993) Protein Engineering 6:327-331).

Anti-ORFX Antibodies

The invention further encompasses antibodies and antibody fragments, such as F_{ab} or (F_{ab})₂, that bind immunospecifically to any of the proteins of the invention.

10 An isolated ORFX protein, or a portion or fragment thereof, can be used as an immunogen to generate antibodies that bind ORFX using standard techniques for polyclonal and monoclonal antibody preparation. Full-length ORFX protein can be used. Alternatively, the invention provides antigenic peptide fragments of ORFX for use as immunogens. The antigenic peptide of ORFX comprises at least 4 amino acid residues of the amino acid sequence shown in
15 any of SEQ ID NO:2*n* (wherein *n* = 1 to 3161). The antigenic peptide encompasses an epitope of ORFX such that an antibody raised against the peptide forms a specific immune complex with ORFX. The antigenic peptide may comprise at least 6 aa residues, at least 8 aa residues, at least 10 aa residues, at least 15 aa residues, at least 20 aa residues, or at least 30 aa residues. In one embodiment of the invention, the antigenic peptide comprises a polypeptide comprising at least 6
20 contiguous amino acids of any of SEQ ID NO:2*n* (wherein *n* = 1 to 3161).

In an embodiment of the invention, epitopes encompassed by the antigenic peptide are regions of ORFX that are located on the surface of the protein, *e.g.*, hydrophilic regions. As a means for targeting antibody production, hydropathy plots showing regions of hydrophilicity and hydrophobicity may be generated by any method well known in the art, including, for example,
25 the Kyte Doolittle or the Hopp Woods methods, either with or without Fourier transformation. See, *e.g.*, Hopp and Woods, 1981, Proc. Nat. Acad. Sci. USA 78: 3824-3828; Kyte and Doolittle 1982, J. Mol. Biol. 157: 105-142, each incorporated herein by reference in their entirety.

As disclosed herein, an ORFX protein sequence of any of SEQ ID NO:2*n* (wherein *n* = 1 to 3161), or derivatives, fragments, analogs or homologs thereof, may be utilized as
30 immunogens in the generation of antibodies that immunospecifically-bind these protein components. The term "antibody" as used herein refers to immunoglobulin molecules and

immunologically active portions of immunoglobulin molecules, i.e., molecules that contain an antigen binding site that specifically binds (immunoreacts with) an antigen, such as ORFX. Such antibodies include, but are not limited to, polyclonal, monoclonal, chimeric, single chain, F_{ab} and $F_{(ab)2}$ fragments, and an F_{ab} expression library. In a specific embodiment, antibodies to human ORFX proteins are disclosed. Various procedures known within the art may be used for the production of polyclonal or monoclonal antibodies to a ORFX protein sequence of any of SEQ ID NO:2n (wherein $n = 1$ to 3161) or derivative, fragment, analog or homolog thereof. Some of these proteins are discussed below.

For the production of polyclonal antibodies, various suitable host animals (e.g., rabbit, goat, mouse or other mammal) may be immunized by injection with the native protein, or a synthetic variant thereof, or a derivative of the foregoing. An appropriate immunogenic preparation can contain, for example, recombinantly expressed ORFX protein or a chemically synthesized ORFX polypeptide. The preparation can further include an adjuvant. Various adjuvants used to increase the immunological response include, but are not limited to, Freund's (complete and incomplete), mineral gels (e.g., aluminum hydroxide), surface active substances (e.g., lysolecithin, pluronic polyols, polyanions, peptides, oil emulsions, dinitrophenol, etc.), human adjuvants such as *Bacille Calmette-Guerin* and *Corynebacterium parvum*, or similar immunostimulatory agents. If desired, the antibody molecules directed against ORFX can be isolated from the mammal (e.g., from the blood) and further purified by well known techniques, such as protein A chromatography to obtain the IgG fraction.

The term "monoclonal antibody" or "monoclonal antibody composition", as used herein, refers to a population of antibody molecules that contain only one species of an antigen binding site capable of immunoreacting with a particular epitope of ORFX. A monoclonal antibody composition thus typically displays a single binding affinity for a particular ORFX protein with which it immunoreacts. For preparation of monoclonal antibodies directed towards a particular ORFX protein, or derivatives, fragments, analogs or homologs thereof, any technique that provides for the production of antibody molecules by continuous cell line culture may be utilized. Such techniques include, but are not limited to, the hybridoma technique (see Kohler & Milstein, 1975 *Nature* 256: 495-497); the trioma technique; the human B-cell hybridoma technique (see Kozbor, *et al.*, 1983 *Immunol Today* 4: 72) and the EBV hybridoma technique to produce human monoclonal antibodies (see Cole, *et al.*, 1985 In: MONOCLONAL ANTIBODIES AND

CANCER THERAPY, Alan R. Liss, Inc., pp. 77-96). Human monoclonal antibodies may be utilized in the practice of the present invention and may be produced by using human hybridomas (see Cote, *et al.*, 1983. *Proc Natl Acad Sci USA* 80: 2026-2030) or by transforming human B-cells with Epstein Barr Virus *in vitro* (see Cole, *et al.*, 1985 In: MONOCLONAL ANTIBODIES AND
5 CANCER THERAPY, Alan R. Liss, Inc., pp. 77-96). Each of the above citations are incorporated herein by reference in their entirety

According to the invention, techniques can be adapted for the production of single-chain antibodies specific to a ORFX protein (see *e.g.*, U.S. Patent No. 4,946,778). In addition, methods can be adapted for the construction of F_{ab} expression libraries (see *e.g.*, Huse, *et al.*,
10 1989 *Science* 246: 1275-1281) to allow rapid and effective identification of monoclonal F_{ab} fragments with the desired specificity for a ORFX protein or derivatives, fragments, analogs or homologs thereof. Non-human antibodies can be "humanized" by techniques well known in the art. See *e.g.*, U.S. Patent No. 5,225,539. Each of the above citations are incorporated herein by reference. Antibody fragments that contain the idiotype to a ORFX protein may be produced by
15 techniques known in the art including, but not limited to: (i) an F_{(ab')₂} fragment produced by pepsin digestion of an antibody molecule; (ii) an F_{ab} fragment generated by reducing the disulfide bridges of an F_{(ab')₂} fragment; (iii) an F_{ab} fragment generated by the treatment of the antibody molecule with papain and a reducing agent and (iv) F_v fragments.

Additionally, recombinant anti-ORFX antibodies, such as chimeric and humanized
20 monoclonal antibodies, comprising both human and non-human portions, which can be made using standard recombinant DNA techniques, are within the scope of the invention. Such chimeric and humanized monoclonal antibodies can be produced by recombinant DNA techniques known in the art, for example using methods described in PCT International Application No. PCT/US86/02269; European Patent Application No. 184,187; European Patent
25 Application No. 171,496; European Patent Application No. 173,494; PCT International Publication No. WO 86/01533; U.S. Pat. No. 4,816,567; European Patent Application No. 125,023; Better *et al.* (1988) *Science* 240:1041-1043; Liu *et al.* (1987) *PNAS* 84:3439-3443; Liu *et al.* (1987) *J Immunol.* 139:3521-3526; Sun *et al.* (1987) *PNAS* 84:214-218; Nishimura *et al.* (1987) *Cancer Res* 47:999-1005; Wood *et al.* (1985) *Nature* 314:446-449; Shaw *et al.* (1988), *J. Natl Cancer Inst* 80:1553-1559; Morrison (1985) *Science* 229:1202-1207; Oi *et al.* (1986) *BioTechniques* 4:214; U.S. Pat. No. 5,225,539; Jones *et al.* (1986) *Nature* 321:552-525;

Verhoeyan *et al.* (1988) *Science* 239:1534; and Beidler *et al.* (1988) *J Immunol* 141:4053-4060. Each of the above citations are incorporated herein by reference.

In one embodiment, methods for the screening of antibodies that possess the desired specificity include, but are not limited to, enzyme-linked immunosorbent assay (ELISA) and other immunologically-mediated techniques known within the art. In a specific embodiment, selection of antibodies that are specific to a particular domain of a ORFX protein is facilitated by generation of hybridomas that bind to the fragment of a ORFX protein possessing such a domain. Antibodies that are specific for one or more domains within a ORFX protein, *e.g.*, the domain spanning the first fifty amino-terminal residues specific to ORFX when compared to FGF-9, or derivatives, fragments, analogs or homologs thereof, are also provided herein.

Anti-ORFX antibodies may be used in methods known within the art relating to the localization and/or quantitation of a ORFX protein (*e.g.*, for use in measuring levels of the ORFX protein within appropriate physiological samples, for use in diagnostic methods, for use in imaging the protein, and the like). In a given embodiment, antibodies for ORFX proteins, or derivatives, fragments, analogs or homologs thereof, that contain the antibody derived binding domain, are utilized as pharmacologically-active compounds [hereinafter "Therapeutics"].

An anti-ORFX antibody (*e.g.*, monoclonal antibody) can be used to isolate ORFX by standard techniques, such as affinity chromatography or immunoprecipitation. An anti-ORFX antibody can facilitate the purification of natural ORFX from cells and of recombinantly produced ORFX expressed in host cells. Moreover, an anti-ORFX antibody can be used to detect ORFX protein (*e.g.*, in a cellular lysate or cell supernatant) in order to evaluate the abundance and pattern of expression of the ORFX protein. Anti-ORFX antibodies can be used diagnostically to monitor protein levels in tissue as part of a clinical testing procedure, *e.g.*, to, for example, determine the efficacy of a given treatment regimen. Detection can be facilitated by coupling (*i.e.*, physically linking) the antibody to a detectable substance. Examples of detectable substances include various enzymes, prosthetic groups, fluorescent materials, luminescent materials, bioluminescent materials, and radioactive materials. Examples of suitable enzymes include horseradish peroxidase, alkaline phosphatase, β -galactosidase, or acetylcholinesterase; examples of suitable prosthetic group complexes include streptavidin/biotin and avidin/biotin; examples of suitable fluorescent materials include umbelliferone, fluorescein, fluorescein isothiocyanate, rhodamine, dichlorotriazinylamine fluorescein, dansyl chloride or

phycoerythrin; an example of a luminescent material includes luminol; examples of bioluminescent materials include luciferase, luciferin, and aequorin, and examples of suitable radioactive material include ^{125}I , ^{131}I , ^{35}S or ^3H .

ORFX Recombinant Vectors and Host Cells

5 Another aspect of the invention pertains to vectors, preferably expression vectors, containing a nucleic acid encoding ORFX protein, or derivatives, fragments, analogs or homologs thereof. As used herein, the term "vector" refers to a nucleic acid molecule capable of transporting another nucleic acid to which it has been linked. One type of vector is a "plasmid", which refers to a circular double stranded DNA loop into which additional DNA segments can
10 be ligated. Another type of vector is a viral vector, wherein additional DNA segments can be ligated into the viral genome. Certain vectors are capable of autonomous replication in a host cell into which they are introduced (*e.g.*, bacterial vectors having a bacterial origin of replication and episomal mammalian vectors). Other vectors (*e.g.*, non-episomal mammalian vectors) are integrated into the genome of a host cell upon introduction into the host cell, and thereby are
15 replicated along with the host genome. Moreover, certain vectors are capable of directing the expression of genes to which they are operatively linked. Such vectors are referred to herein as "expression vectors". In general, expression vectors of utility in recombinant DNA techniques are often in the form of plasmids. In the present specification, "plasmid" and "vector" can be used interchangeably as the plasmid is the most commonly used form of vector. However, the
20 invention is intended to include such other forms of expression vectors, such as viral vectors (*e.g.*, replication defective retroviruses, adenoviruses and adeno-associated viruses), which serve equivalent functions.

The recombinant expression vectors of the invention comprise a nucleic acid of the invention in a form suitable for expression of the nucleic acid in a host cell, which means that the
25 recombinant expression vectors include one or more regulatory sequences, selected on the basis of the host cells to be used for expression, that is operatively linked to the nucleic acid sequence to be expressed. Within a recombinant expression vector, "operably linked" is intended to mean that the nucleotide sequence of interest is linked to the regulatory sequence(s) in a manner that allows for expression of the nucleotide sequence (*e.g.*, in an *in vitro* transcription/translation
30 system or in a host cell when the vector is introduced into the host cell). The term "regulatory sequence" is intended to include promoters, enhancers and other expression control elements

(e.g., polyadenylation signals). Such regulatory sequences are described, for example, in Goeddel; GENE EXPRESSION TECHNOLOGY: METHODS IN ENZYMOLOGY 185, Academic Press, San Diego, Calif. (1990). Regulatory sequences include those that direct constitutive expression of a nucleotide sequence in many types of host cell and those that direct expression of the
5 nucleotide sequence only in certain host cells (e.g., tissue-specific regulatory sequences). It will be appreciated by those skilled in the art that the design of the expression vector can depend on such factors as the choice of the host cell to be transformed, the level of expression of protein desired, etc. The expression vectors of the invention can be introduced into host cells to thereby produce proteins or peptides, including fusion proteins or peptides, encoded by nucleic acids as
10 described herein (e.g., ORFX proteins, mutant forms of ORFX, fusion proteins, etc.).

The recombinant expression vectors of the invention can be designed for expression of ORFX in prokaryotic or eukaryotic cells. For example, ORFX can be expressed in bacterial cells such as *E. coli*, insect cells (using baculovirus expression vectors) yeast cells or mammalian cells. Suitable host cells are discussed further in Goeddel, GENE EXPRESSION TECHNOLOGY:
15 METHODS IN ENZYMOLOGY 185, Academic Press, San Diego, Calif. (1990). Alternatively, the recombinant expression vector can be transcribed and translated *in vitro*, for example using T7 promoter regulatory sequences and T7 polymerase.

Expression of proteins in prokaryotes is most often carried out in *E. coli* with vectors containing constitutive or inducible promoters directing the expression of either fusion or
20 non-fusion proteins. Fusion vectors add a number of amino acids to a protein encoded therein, usually to the amino terminus of the recombinant protein. Such fusion vectors typically serve three purposes: (1) to increase expression of recombinant protein; (2) to increase the solubility of the recombinant protein; and (3) to aid in the purification of the recombinant protein by acting as a ligand in affinity purification. Often, in fusion expression vectors, a proteolytic cleavage site is
25 introduced at the junction of the fusion moiety and the recombinant protein to enable separation of the recombinant protein from the fusion moiety subsequent to purification of the fusion protein. Such enzymes, and their cognate recognition sequences, include Factor Xa, thrombin and enterokinase. Typical fusion expression vectors include pGEX (Pharmacia Biotech Inc; Smith and Johnson (1988) *Gene* 67:31-40), pMAL (New England Biolabs, Beverly, Mass.) and
30 pRIT5 (Pharmacia, Piscataway, N.J.) that fuse glutathione S-transferase (GST), maltose E binding protein, or protein A, respectively, to the target recombinant protein.

Examples of suitable inducible non-fusion *E. coli* expression vectors include pTtrc (Amrann *et al.*, (1988) *Gene* 69:301-315) and pET 11d (Studier *et al.*, GENE EXPRESSION TECHNOLOGY: METHODS IN ENZYMOLOGY 185, Academic Press, San Diego, Calif. (1990) 60-89).

- 5 One strategy to maximize recombinant protein expression in *E. coli* is to express the protein in a host bacteria with an impaired capacity to proteolytically cleave the recombinant protein. See, Gottesman, GENE EXPRESSION TECHNOLOGY: METHODS IN ENZYMOLOGY 185, Academic Press, San Diego, Calif. (1990) 119-128. Another strategy is to alter the nucleic acid sequence of the nucleic acid to be inserted into an expression vector so that the individual codons
10 for each amino acid are those preferentially utilized in *E. coli* (Wada *et al.*, (1992) *Nucleic Acids Res.* 20:2111-2118). Such alteration of nucleic acid sequences of the invention can be carried out by standard DNA synthesis techniques.

- In another embodiment, the ORFX expression vector is a yeast expression vector. Examples of vectors for expression in yeast *S. cerevisiae* include pYepSec1 (Baldari, *et al.*,
15 (1987) *EMBO J* 6:229-234), pMFa (Kurjan and Herskowitz, (1982) *Cell* 30:933-943), pJRY88 (Schultz *et al.*, (1987) *Gene* 54:113-123), pYES2 (Invitrogen Corporation, San Diego, Calif.), and picZ (InVitrogen Corp, San Diego, Calif.).

- Alternatively, ORFX can be expressed in insect cells using baculovirus expression vectors. Baculovirus vectors available for expression of proteins in cultured insect cells (*e.g.*,
20 SF9 cells) include the pAc series (Smith *et al.* (1983) *Mol Cell Biol* 3:2156-2165) and the pVL series (Lucklow and Summers (1989) *Virology* 170:31-39).

- In yet another embodiment, a nucleic acid of the invention is expressed in mammalian cells using a mammalian expression vector. Examples of mammalian expression vectors include pCDM8 (Seed (1987) *Nature* 329:840) and pMT2PC (Kaufman *et al.* (1987) *EMBO J*
25 6: 187-195). When used in mammalian cells, the expression vector's control functions are often provided by viral regulatory elements. For example, commonly used promoters are derived from polyoma, Adenovirus 2, cytomegalovirus and Simian Virus 40. For other suitable expression systems for both prokaryotic and eukaryotic cells. See, *e.g.*, Chapters 16 and 17 of Sambrook *et al.*, MOLECULAR CLONING: A LABORATORY MANUAL. 2nd ed., Cold Spring Harbor Laboratory,
30 Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y., 1989.

In another embodiment, the recombinant mammalian expression vector is capable of directing expression of the nucleic acid preferentially in a particular cell type (*e.g.*, tissue-specific regulatory elements are used to express the nucleic acid). Tissue-specific regulatory elements are known in the art. Non-limiting examples of suitable tissue-specific promoters include the albumin promoter (liver-specific; Pinkert *et al.* (1987) *Genes Dev* 1:268-277), lymphoid-specific promoters (Calame and Eaton (1988) *Adv Immunol* 43:235-275), in particular promoters of T cell receptors (Winoto and Baltimore (1989) *EMBO J* 8:729-733) and immunoglobulins (Banerji *et al.* (1983) *Cell* 33:729-740; Queen and Baltimore (1983) *Cell* 33:741-748), neuron-specific promoters (*e.g.*, the neurofilament promoter; Byrne and Ruddle (1989) *PNAS* 86:5473-5477), pancreas-specific promoters (Edlund *et al.* (1985) *Science* 230:912-916), and mammary gland-specific promoters (*e.g.*, milk whey promoter; U.S. Pat. No. 4,873,316 and European Application Publication No. 264,166). Developmentally-regulated promoters are also encompassed, *e.g.*, the murine hox promoters (Kessel and Gruss (1990) *Science* 249:374-379) and the α -fetoprotein promoter (Campes and Tilghman (1989) *Genes Dev* 3:537-546).

The invention further provides a recombinant expression vector comprising a DNA molecule of the invention cloned into the expression vector in an antisense orientation. That is, the DNA molecule is operatively linked to a regulatory sequence in a manner that allows for expression (by transcription of the DNA molecule) of an RNA molecule that is antisense to ORFX mRNA. Regulatory sequences operatively linked to a nucleic acid cloned in the antisense orientation can be chosen that direct the continuous expression of the antisense RNA molecule in a variety of cell types, for instance viral promoters and/or enhancers, or regulatory sequences can be chosen that direct constitutive, tissue specific or cell type specific expression of antisense RNA. The antisense expression vector can be in the form of a recombinant plasmid, phagemid or attenuated virus in which antisense nucleic acids are produced under the control of a high efficiency regulatory region, the activity of which can be determined by the cell type into which the vector is introduced. For a discussion of the regulation of gene expression using antisense genes see Weintraub *et al.*, "Antisense RNA as a molecular tool for genetic analysis," Reviews--Trends in Genetics, Vol. 1(1) 1986.

Another aspect of the invention pertains to host cells into which a recombinant expression vector of the invention has been introduced. The terms "host cell" and "recombinant

host cell" are used interchangeably herein. It is understood that such terms refer not only to the particular subject cell but to the progeny or potential progeny of such a cell. Because certain modifications may occur in succeeding generations due to either mutation or environmental influences, such progeny may not, in fact, be identical to the parent cell, but are still included within the scope of the term as used herein.

A host cell can be any prokaryotic or eukaryotic cell. For example, ORFX protein can be expressed in bacterial cells such as *E. coli*, insect cells, yeast or mammalian cells (such as Chinese hamster ovary cells (CHO) or COS cells). Other suitable host cells are known to those skilled in the art.

Vector DNA can be introduced into prokaryotic or eukaryotic cells via conventional transformation or transfection techniques. As used herein, the terms "transformation" and "transfection" are intended to refer to a variety of art-recognized techniques for introducing foreign nucleic acid (*e.g.*, DNA) into a host cell, including calcium phosphate or calcium chloride co-precipitation, DEAE-dextran-mediated transfection, lipofection, or electroporation. Suitable methods for transforming or transfecting host cells can be found in Sambrook, *et al.* (MOLECULAR CLONING: A LABORATORY MANUAL. 2nd ed., Cold Spring Harbor Laboratory, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y., 1989), and other laboratory manuals.

For stable transfection of mammalian cells, it is known that, depending upon the expression vector and transfection technique used, only a small fraction of cells may integrate the foreign DNA into their genome. In order to identify and select these integrants, a gene that encodes a selectable marker (*e.g.*, resistance to antibiotics) is generally introduced into the host cells along with the gene of interest. Various selectable markers include those that confer resistance to drugs, such as G418, hygromycin and methotrexate. Nucleic acid encoding a selectable marker can be introduced into a host cell on the same vector as that encoding ORFX or can be introduced on a separate vector. Cells stably transfected with the introduced nucleic acid can be identified by drug selection (*e.g.*, cells that have incorporated the selectable marker gene will survive, while the other cells die).

A host cell of the invention, such as a prokaryotic or eukaryotic host cell in culture, can be used to produce (*i.e.*, express) ORFX protein. Accordingly, the invention further provides methods for producing ORFX protein using the host cells of the invention. In one embodiment,

the method comprises culturing the host cell of invention (into which a recombinant expression vector encoding ORFX has been introduced) in a suitable medium such that ORFX protein is produced. In another embodiment, the method further comprises isolating ORFX from the medium or the host cell.

5 **Transgenic animals**

The host cells of the invention can also be used to produce nonhuman transgenic animals. For example, in one embodiment, a host cell of the invention is a fertilized oocyte or an embryonic stem cell into which ORFX-coding sequences have been introduced. Such host cells can then be used to create non-human transgenic animals in which exogenous ORFX sequences
10 have been introduced into their genome or homologous recombinant animals in which endogenous ORFX sequences have been altered. Such animals are useful for studying the function and/or activity of ORFX and for identifying and/or evaluating modulators of ORFX activity. As used herein, a "transgenic animal" is a non-human animal, preferably a mammal, more preferably a rodent such as a rat or mouse, in which one or more of the cells of the animal
15 includes a transgene. Other examples of transgenic animals include non-human primates, sheep, dogs, cows, goats, chickens, amphibians, etc. A transgene is exogenous DNA that is integrated into the genome of a cell from which a transgenic animal develops and that remains in the genome of the mature animal, thereby directing the expression of an encoded gene product in one or more cell types or tissues of the transgenic animal. As used herein, a "homologous
20 recombinant animal" is a non-human animal, preferably a mammal, more preferably a mouse, in which an endogenous ORFX gene has been altered by homologous recombination between the endogenous gene and an exogenous DNA molecule introduced into a cell of the animal, *e.g.*, an embryonic cell of the animal, prior to development of the animal.

A transgenic animal of the invention can be created by introducing ORFX-encoding
25 nucleic acid into the male pronuclei of a fertilized oocyte, *e.g.*, by microinjection, retroviral infection, and allowing the oocyte to develop in a pseudopregnant female foster animal. The human ORFX DNA sequence of SEQ ID NO:2*n*-1 (wherein *n* = 1 to 3161) can be introduced as a transgene into the genome of a non-human animal. Alternatively, a nonhuman homologue of the human ORFX gene, such as a mouse ORFX gene, can be isolated based on hybridization to
30 the human ORFX cDNA (described further above) and used as a transgene. Intronic sequences and polyadenylation signals can also be included in the transgene to increase the efficiency of

expression of the transgene. A tissue-specific regulatory sequence(s) can be operably linked to the ORFX transgene to direct expression of ORFX protein to particular cells. Methods for generating transgenic animals via embryo manipulation and microinjection, particularly animals such as mice, have become conventional in the art and are described, for example, in U.S. Pat. Nos. 4,736,866; 4,870,009; and 4,873,191; and Hogan 1986, In: MANIPULATING THE MOUSE EMBRYO, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y. Similar methods are used for production of other transgenic animals. A transgenic founder animal can be identified based upon the presence of the ORFX transgene in its genome and/or expression of ORFX mRNA in tissues or cells of the animals. A transgenic founder animal can then be used to breed additional animals carrying the transgene. Moreover, transgenic animals carrying a transgene encoding ORFX can further be bred to other transgenic animals carrying other transgenes.

To create a homologous recombinant animal, a vector is prepared which contains at least a portion of a ORFX gene into which a deletion, addition or substitution has been introduced to thereby alter, *e.g.*, functionally disrupt, the ORFX gene. The ORFX gene can be a human gene (*e.g.*, SEQ ID NO:2*n*-1 (wherein *n* = 1 to 3161)), but more preferably, is a non-human homologue of a human ORFX gene. For example, a mouse homologue of human ORFX gene of SEQ ID NO:2*n*-1 (wherein *n* = 1 to 3161) can be used to construct a homologous recombination vector suitable for altering an endogenous ORFX gene in the mouse genome. In one embodiment, the vector is designed such that, upon homologous recombination, the endogenous ORFX gene is functionally disrupted (*i.e.*, no longer encodes a functional protein; also referred to as a "knock out" vector).

Alternatively, the vector can be designed such that, upon homologous recombination, the endogenous ORFX gene is mutated or otherwise altered but still encodes functional protein (*e.g.*, the upstream regulatory region can be altered to thereby alter the expression of the endogenous ORFX protein). In the homologous recombination vector, the altered portion of the ORFX gene is flanked at its 5' and 3' ends by additional nucleic acid of the ORFX gene to allow for homologous recombination to occur between the exogenous ORFX gene carried by the vector and an endogenous ORFX gene in an embryonic stem cell. The additional flanking ORFX nucleic acid is of sufficient length for successful homologous recombination with the endogenous gene. Typically, several kilobases of flanking DNA (both at the 5' and 3' ends) are included in the vector. See *e.g.*, Thomas *et al.* (1987) *Cell* 51:503 for a description of

homologous recombination vectors. The vector is introduced into an embryonic stem cell line (e.g., by electroporation) and cells in which the introduced ORFX gene has homologously recombined with the endogenous ORFX gene are selected (see e.g., Li *et al.* (1992) *Cell* 69:915).

5 The selected cells are then injected into a blastocyst of an animal (e.g., a mouse) to form aggregation chimeras. See e.g., Bradley 1987, In: TERATOCARCINOMAS AND EMBRYONIC STEM CELLS: A PRACTICAL APPROACH, Robertson, ed. IRL, Oxford, pp. 113-152. A chimeric embryo can then be implanted into a suitable pseudopregnant female foster animal and the embryo brought to term. Progeny harboring the homologously recombined DNA in their germ cells can be used to breed animals in which all cells of the animal contain the homologously recombined
10 DNA by germline transmission of the transgene. Methods for constructing homologous recombination vectors and homologous recombinant animals are described further in Bradley (1991) *Curr Opin Biotechnol* 2:823-829; PCT International Publication Nos.: WO 90/11354; WO 91/01140; WO 92/0968; and WO 93/04169.

In another embodiment, transgenic non-humans animals can be produced that contain
15 selected systems that allow for regulated expression of the transgene. One example of such a system is the cre/loxP recombinase system of bacteriophage P1. For a description of the cre/loxP recombinase system, see, e.g., Lakso *et al.* (1992) *PNAS* 89:6232-6236. Another example of a recombinase system is the FLP recombinase system of *Saccharomyces cerevisiae* (O'Gorman *et al.* (1991) *Science* 251:1351-1355. If a cre/loxP recombinase system is used to regulate
20 expression of the transgene, animals containing transgenes encoding both the Cre recombinase and a selected protein are required. Such animals can be provided through the construction of "double" transgenic animals, e.g., by mating two transgenic animals, one containing a transgene encoding a selected protein and the other containing a transgene encoding a recombinase.

Clones of the non-human transgenic animals described herein can also be produced
25 according to the methods described in Wilmot *et al.* (1997) *Nature* 385:810-813. In brief, a cell, e.g., a somatic cell, from the transgenic animal can be isolated and induced to exit the growth cycle and enter G₀ phase. The quiescent cell can then be fused, e.g., through the use of electrical pulses, to an enucleated oocyte from an animal of the same species from which the quiescent cell is isolated. The reconstructed oocyte is then cultured such that it develops to morula or
30 blastocyte and then transferred to pseudopregnant female foster animal. The offspring borne of

this female foster animal will be a clone of the animal from which the cell, *e.g.*, the somatic cell, is isolated.

Pharmaceutical Compositions

5 The ORFX nucleic acid molecules, ORFX proteins, and anti-ORFX antibodies (also referred to herein as "active compounds") of the invention, and derivatives, fragments, analogs and homologs thereof, can be incorporated into pharmaceutical compositions suitable for administration. Such compositions typically comprise the nucleic acid molecule, protein, or antibody and a pharmaceutically acceptable carrier. As used herein, "pharmaceutically acceptable carrier" is intended to include any and all solvents, dispersion media, coatings, 10 antibacterial and antifungal agents, isotonic and absorption delaying agents, and the like, compatible with pharmaceutical administration. Suitable carriers are described in the most recent edition of Remington's Pharmaceutical Sciences, a standard reference text in the field, which is incorporated herein by reference. Preferred examples of such carriers or diluents include, but are not limited to, water, saline, finger's solutions, dextrose solution, and 5% human 15 serum albumin. Liposomes and non-aqueous vehicles such as fixed oils may also be used. The use of such media and agents for pharmaceutically active substances is well known in the art. Except insofar as any conventional media or agent is incompatible with the active compound, use thereof in the compositions is contemplated. Supplementary active compounds can also be incorporated into the compositions.

20 A pharmaceutical composition of the invention is formulated to be compatible with its intended route of administration. Examples of routes of administration include parenteral, *e.g.*, intravenous, intradermal, subcutaneous, oral (*e.g.*, inhalation), transdermal (topical), transmucosal, and rectal administration. Solutions or suspensions used for parenteral, intradermal, or subcutaneous application can include the following components: a sterile diluent 25 such as water for injection, saline solution, fixed oils, polyethylene glycols, glycerine, propylene glycol or other synthetic solvents; antibacterial agents such as benzyl alcohol or methyl parabens; antioxidants such as ascorbic acid or sodium bisulfite; chelating agents such as ethylenediaminetetraacetic acid; buffers such as acetates, citrates or phosphates, and agents for the adjustment of tonicity such as sodium chloride or dextrose. The pH can be adjusted with 30 acids or bases, such as hydrochloric acid or sodium hydroxide. The parenteral preparation can be enclosed in ampoules, disposable syringes or multiple dose vials made of glass or plastic.

Pharmaceutical compositions suitable for injectable use include sterile aqueous solutions (where water soluble) or dispersions and sterile powders for the extemporaneous preparation of sterile injectable solutions or dispersion. For intravenous administration, suitable carriers include physiological saline, bacteriostatic water, Cremophor EL™ (BASF, Parsippany, N.J.) or phosphate buffered saline (PBS). In all cases, the composition must be sterile and should be fluid to the extent that easy syringeability exists. It must be stable under the conditions of manufacture and storage and must be preserved against the contaminating action of microorganisms such as bacteria and fungi. The carrier can be a solvent or dispersion medium containing, for example, water, ethanol, polyol (for example, glycerol, propylene glycol, and liquid polyethylene glycol, and the like), and suitable mixtures thereof. The proper fluidity can be maintained, for example, by the use of a coating such as lecithin, by the maintenance of the required particle size in the case of dispersion and by the use of surfactants. Prevention of the action of microorganisms can be achieved by various antibacterial and antifungal agents, for example, parabens, chlorobutanol, phenol, ascorbic acid, thimerosal, and the like. In many cases, it will be preferable to include isotonic agents, for example, sugars, polyalcohols such as manitol, sorbitol, sodium chloride in the composition. Prolonged absorption of the injectable compositions can be brought about by including in the composition an agent which delays absorption, for example, aluminum monostearate and gelatin.

Sterile injectable solutions can be prepared by incorporating the active compound (e.g., a ORFX protein or anti-ORFX antibody) in the required amount in an appropriate solvent with one or a combination of ingredients enumerated above, as required, followed by filtered sterilization. Generally, dispersions are prepared by incorporating the active compound into a sterile vehicle that contains a basic dispersion medium and the required other ingredients from those enumerated above. In the case of sterile powders for the preparation of sterile injectable solutions, methods of preparation are vacuum drying and freeze-drying that yields a powder of the active ingredient plus any additional desired ingredient from a previously sterile-filtered solution thereof.

Oral compositions generally include an inert diluent or an edible carrier. They can be enclosed in gelatin capsules or compressed into tablets. For the purpose of oral therapeutic administration, the active compound can be incorporated with excipients and used in the form of tablets, troches, or capsules. Oral compositions can also be prepared using a fluid carrier for use

as a mouthwash, wherein the compound in the fluid carrier is applied orally and swished and expectorated or swallowed. Pharmaceutically compatible binding agents, and/or adjuvant materials can be included as part of the composition. The tablets, pills, capsules, troches and the like can contain any of the following ingredients, or compounds of a similar nature: a binder such as microcrystalline cellulose, gum tragacanth or gelatin; an excipient such as starch or lactose, a disintegrating agent such as alginic acid, Primogel, or corn starch; a lubricant such as magnesium stearate or Sterotes; a glidant such as colloidal silicon dioxide; a sweetening agent such as sucrose or saccharin; or a flavoring agent such as peppermint, methyl salicylate, or orange flavoring.

For administration by inhalation, the compounds are delivered in the form of an aerosol spray from pressured container or dispenser which contains a suitable propellant, *e.g.*, a gas such as carbon dioxide, or a nebulizer.

Systemic administration can also be by transmucosal or transdermal means. For transmucosal or transdermal administration, penetrants appropriate to the barrier to be permeated are used in the formulation. Such penetrants are generally known in the art, and include, for example, for transmucosal administration, detergents, bile salts, and fusidic acid derivatives. Transmucosal administration can be accomplished through the use of nasal sprays or suppositories. For transdermal administration, the active compounds are formulated into ointments, salves, gels, or creams as generally known in the art.

The compounds can also be prepared in the form of suppositories (*e.g.*, with conventional suppository bases such as cocoa butter and other glycerides) or retention enemas for rectal delivery.

In one embodiment, the active compounds are prepared with carriers that will protect the compound against rapid elimination from the body, such as a controlled release formulation, including implants and microencapsulated delivery systems. Biodegradable, biocompatible polymers can be used, such as ethylene vinyl acetate, polyanhydrides, polyglycolic acid, collagen, polyorthoesters, and polylactic acid. Methods for preparation of such formulations will be apparent to those skilled in the art. The materials can also be obtained commercially from Alza Corporation and Nova Pharmaceuticals, Inc. Liposomal suspensions (including liposomes targeted to infected cells with monoclonal antibodies to viral antigens) can also be used as pharmaceutically acceptable carriers. These can be prepared according to methods known to those skilled in the art, for example, as described in U.S. Pat. No. 4,522,811.

It is especially advantageous to formulate oral or parenteral compositions in dosage unit form for ease of administration and uniformity of dosage. Dosage unit form as used herein refers to physically discrete units suited as unitary dosages for the subject to be treated; each unit containing a predetermined quantity of active compound calculated to produce the desired
5 therapeutic effect in association with the required pharmaceutical carrier. The specification for the dosage unit forms of the invention are dictated by and directly dependent on the unique characteristics of the active compound and the particular therapeutic effect to be achieved.

The nucleic acid molecules of the invention can be inserted into vectors and used as gene therapy vectors. Gene therapy vectors can be delivered to a subject by any of a number of
10 routes, *e.g.*, as described in U.S. Patent Nos. 5,703,055. Delivery can thus also include, *e.g.*, intravenous injection, local administration (see U.S. Pat. No. 5,328,470) or stereotactic injection (see *e.g.*, Chen *et al.* (1994) *PNAS* 91:3054-3057). The pharmaceutical preparation of the gene therapy vector can include the gene therapy vector in an acceptable diluent, or can comprise a slow release matrix in which the gene delivery vehicle is imbedded. Alternatively, where the
15 complete gene delivery vector can be produced intact from recombinant cells, *e.g.*, retroviral vectors, the pharmaceutical preparation can include one or more cells that produce the gene delivery system.

The pharmaceutical compositions can be included in a container, pack, or dispenser together with instructions for administration.

20 Additional Uses and Methods of the Invention

The nucleic acid molecules, proteins, protein homologues, and antibodies described herein can be used in one or more of the following methods: (a) screening assays; (b) detection assays (*e.g.*, chromosomal mapping, cell and tissue typing, forensic biology), (c) predictive
25 medicine (*e.g.*, diagnostic assays, prognostic assays, monitoring clinical trials, and pharmacogenomics); and (d) methods of treatment (*e.g.*, therapeutic and prophylactic).

The isolated nucleic acid molecules of the invention can be used to express ORFX protein (*e.g.*, via a recombinant expression vector in a host cell in gene therapy applications), to detect ORFX mRNA (*e.g.*, in a biological sample) or a genetic lesion in a ORFX gene, and to modulate ORFX activity, as described further below. In addition, the ORFX proteins can be
30 used to screen drugs or compounds that modulate the ORFX activity or expression as well as to treat disorders characterized by insufficient or excessive production of ORFX protein, for

example proliferative or differentiative disorders, or production of ORFX protein forms that have decreased or aberrant activity compared to ORFX wild type protein. In addition, the anti-ORFX antibodies of the invention can be used to detect and isolate ORFX proteins and modulate ORFX activity.

- 5 This invention further pertains to novel agents identified by the above described screening assays and uses thereof for treatments as described herein.

Screening Assays

- The invention provides a method (also referred to herein as a "screening assay") for identifying modulators, *i.e.*, candidate or test compounds or agents (*e.g.*, peptides,
10 peptidomimetics, small molecules or other drugs) that bind to ORFX proteins or have a stimulatory or inhibitory effect on, for example, ORFX expression or ORFX activity.

- In one embodiment, the invention provides assays for screening candidate or test compounds which bind to or modulate the activity of a ORFX protein or polypeptide or biologically active portion thereof. The test compounds of the present invention can be obtained
15 using any of the numerous approaches in combinatorial library methods known in the art, including: biological libraries; spatially addressable parallel solid phase or solution phase libraries; synthetic library methods requiring deconvolution; the "one-bead one-compound" library method; and synthetic library methods using affinity chromatography selection. The biological library approach is limited to peptide libraries, while the other four approaches are
20 applicable to peptide, non-peptide oligomer or small molecule libraries of compounds (Lam (1997) *Anticancer Drug Des* 12:145).

- Examples of methods for the synthesis of molecular libraries can be found in the art, for example in: DeWitt *et al.* (1993) *Proc Natl Acad Sci U.S.A.* 90:6909; Erb *et al.* (1994) *Proc Natl Acad Sci U.S.A.* 91:11422; Zuckermann *et al.* (1994) *J Med Chem* 37:2678; Cho *et al.* (1993)
25 *Science* 261:1303; Carrell *et al.* (1994) *Angew Chem Int Ed Engl* 33:2059; Carell *et al.* (1994) *Angew Chem Int Ed Engl* 33:2061; and Gallop *et al.* (1994) *J Med Chem* 37:1233.

- Libraries of compounds may be presented in solution (*e.g.*, Houghten (1992) *Biotechniques* 13:412-421), or on beads (Lam (1991) *Nature* 354:82-84), on chips (Fodor (1993) *Nature* 364:555-556), bacteria (Ladner U.S. Pat. No. 5,223,409), spores (Ladner USP '409),
30 plasmids (Cull *et al.* (1992) *Proc Natl Acad Sci USA* 89:1865-1869) or on phage (Scott and

Smith (1990) *Science* 249:386-390; Devlin (1990) *Science* 249:404-406; Cwirla *et al.* (1990) *Proc Natl Acad Sci U.S.A.* 87:6378-6382; Felici (1991) *J Mol Biol* 222:301-310; Ladner *above.*).

In one embodiment, an assay is a cell-based assay in which a cell which expresses a membrane-bound form of ORFX protein, or a biologically active portion thereof, on the cell surface is contacted with a test compound and the ability of the test compound to bind to a ORFX protein determined. The cell, for example, can of mammalian origin or a yeast cell. Determining the ability of the test compound to bind to the ORFX protein can be accomplished, for example, by coupling the test compound with a radioisotope or enzymatic label such that binding of the test compound to the ORFX protein or biologically active portion thereof can be determined by detecting the labeled compound in a complex. For example, test compounds can be labeled with ^{125}I , ^{35}S , ^{14}C , or ^3H , either directly or indirectly, and the radioisotope detected by direct counting of radioemission or by scintillation counting. Alternatively, test compounds can be enzymatically labeled with, for example, horseradish peroxidase, alkaline phosphatase, or luciferase, and the enzymatic label detected by determination of conversion of an appropriate substrate to product. In one embodiment, the assay comprises contacting a cell which expresses a membrane-bound form of ORFX protein, or a biologically active portion thereof, on the cell surface with a known compound which binds ORFX to form an assay mixture, contacting the assay mixture with a test compound, and determining the ability of the test compound to interact with a ORFX protein, wherein determining the ability of the test compound to interact with a ORFX protein comprises determining the ability of the test compound to preferentially bind to ORFX or a biologically active portion thereof as compared to the known compound.

In another embodiment, an assay is a cell-based assay comprising contacting a cell expressing a membrane-bound form of ORFX protein, or a biologically active portion thereof, on the cell surface with a test compound and determining the ability of the test compound to modulate (*e.g.*, stimulate or inhibit) the activity of the ORFX protein or biologically active portion thereof. Determining the ability of the test compound to modulate the activity of ORFX or a biologically active portion thereof can be accomplished, for example, by determining the ability of the ORFX protein to bind to or interact with a ORFX target molecule. As used herein, a "target molecule" is a molecule with which a ORFX protein binds or interacts in nature, for example, a molecule on the surface of a cell which expresses a ORFX interacting protein, a molecule on the surface of a second cell, a molecule in the extracellular milieu, a molecule

associated with the internal surface of a cell membrane or a cytoplasmic molecule. A ORFX target molecule can be a non-ORFX molecule or a ORFX protein or polypeptide of the present invention. In one embodiment, a ORFX target molecule is a component of a signal transduction pathway that facilitates transduction of an extracellular signal (*e.g.*, a signal generated by binding of a compound to a membrane-bound ORFX molecule) through the cell membrane and into the cell. The target, for example, can be a second intercellular protein that has catalytic activity or a protein that facilitates the association of downstream signaling molecules with ORFX.

Determining the ability of the ORFX protein to bind to or interact with a ORFX target molecule can be accomplished by one of the methods described above for determining direct binding. In one embodiment, determining the ability of the ORFX protein to bind to or interact with a ORFX target molecule can be accomplished by determining the activity of the target molecule. For example, the activity of the target molecule can be determined by detecting induction of a cellular second messenger of the target (*i.e.* intracellular Ca^{2+} , diacylglycerol, IP_3 , etc.), detecting catalytic/enzymatic activity of the target an appropriate substrate, detecting the induction of a reporter gene (comprising a ORFX-responsive regulatory element operatively linked to a nucleic acid encoding a detectable marker, *e.g.*, luciferase), or detecting a cellular response, for example, cell survival, cellular differentiation, or cell proliferation.

In yet another embodiment, an assay of the present invention is a cell-free assay comprising contacting a ORFX protein or biologically active portion thereof with a test compound and determining the ability of the test compound to bind to the ORFX protein or biologically active portion thereof. Binding of the test compound to the ORFX protein can be determined either directly or indirectly as described above. In one embodiment, the assay comprises contacting the ORFX protein or biologically active portion thereof with a known compound which binds ORFX to form an assay mixture, contacting the assay mixture with a test compound, and determining the ability of the test compound to interact with a ORFX protein, wherein determining the ability of the test compound to interact with a ORFX protein comprises determining the ability of the test compound to preferentially bind to ORFX or biologically active portion thereof as compared to the known compound.

In another embodiment, an assay is a cell-free assay comprising contacting ORFX protein or biologically active portion thereof with a test compound and determining the ability of the test compound to modulate (*e.g.*, stimulate or inhibit) the activity of the ORFX protein or

biologically active portion thereof. Determining the ability of the test compound to modulate the activity of ORFX can be accomplished, for example, by determining the ability of the ORFX protein to bind to a ORFX target molecule by one of the methods described above for determining direct binding. In an alternative embodiment, determining the ability of the test compound to modulate the activity of ORFX can be accomplished by determining the ability of the ORFX protein further modulate a ORFX target molecule. For example, the catalytic/enzymatic activity of the target molecule on an appropriate substrate can be determined as previously described.

In yet another embodiment, the cell-free assay comprises contacting the ORFX protein or biologically active portion thereof with a known compound which binds ORFX to form an assay mixture, contacting the assay mixture with a test compound, and determining the ability of the test compound to interact with a ORFX protein, wherein determining the ability of the test compound to interact with a ORFX protein comprises determining the ability of the ORFX protein to preferentially bind to or modulate the activity of a ORFX target molecule.

The cell-free assays of the present invention are amenable to use of both the soluble form or the membrane-bound form of ORFX. In the case of cell-free assays comprising the membrane-bound form of ORFX, it may be desirable to utilize a solubilizing agent such that the membrane-bound form of ORFX is maintained in solution. Examples of such solubilizing agents include non-ionic detergents such as n-octylglucoside, n-dodecylglucoside, n-dodecylmaltoside, octanoyl-N-methylglucamide, decanoyl-N-methylglucamide, Triton® X-100, Triton® X-114, Thesit®, Isotridecypoly(ethylene glycol ether)_n, N-dodecyl--N,N-dimethyl-3-ammonio-1-propane sulfonate, 3-(3-cholamidopropyl)dimethylamminiol-1-propane sulfonate (CHAPS), or 3-(3-cholamidopropyl)dimethylamminiol-2-hydroxy-1-propane sulfonate (CHAPSO).

In more than one embodiment of the above assay methods of the present invention, it may be desirable to immobilize either ORFX or its target molecule to facilitate separation of complexed from uncomplexed forms of one or both of the proteins, as well as to accommodate automation of the assay. Binding of a test compound to ORFX, or interaction of ORFX with a target molecule in the presence and absence of a candidate compound, can be accomplished in any vessel suitable for containing the reactants. Examples of such vessels include microtiter plates, test tubes, and micro-centrifuge tubes. In one embodiment, a fusion protein can be provided that adds a domain that allows one or both of the proteins to be bound to a matrix. For

example, GST-ORFX fusion proteins or GST-target fusion proteins can be adsorbed onto glutathione sepharose beads (Sigma Chemical, St. Louis, MO) or glutathione derivatized microtiter plates, that are then combined with the test compound or the test compound and either the non-adsorbed target protein or ORFX protein, and the mixture is incubated under conditions conducive to complex formation (e.g., at physiological conditions for salt and pH). Following incubation, the beads or microtiter plate wells are washed to remove any unbound components, the matrix immobilized in the case of beads, complex determined either directly or indirectly, for example, as described above. Alternatively, the complexes can be dissociated from the matrix, and the level of ORFX binding or activity determined using standard techniques.

Other techniques for immobilizing proteins on matrices can also be used in the screening assays of the invention. For example, either ORFX or its target molecule can be immobilized utilizing conjugation of biotin and streptavidin. Biotinylated ORFX or target molecules can be prepared from biotin-NHS (N-hydroxy-succinimide) using techniques well known in the art (e.g., biotinylation kit, Pierce Chemicals, Rockford, Ill.), and immobilized in the wells of streptavidin-coated 96 well plates (Pierce Chemical). Alternatively, antibodies reactive with ORFX or target molecules, but which do not interfere with binding of the ORFX protein to its target molecule, can be derivatized to the wells of the plate, and unbound target or ORFX trapped in the wells by antibody conjugation. Methods for detecting such complexes, in addition to those described above for the GST-immobilized complexes, include immunodetection of complexes using antibodies reactive with the ORFX or target molecule, as well as enzyme-linked assays that rely on detecting an enzymatic activity associated with the ORFX or target molecule.

In another embodiment, modulators of ORFX expression are identified in a method wherein a cell is contacted with a candidate compound and the expression of ORFX mRNA or protein in the cell is determined. The level of expression of ORFX mRNA or protein in the presence of the candidate compound is compared to the level of expression of ORFX mRNA or protein in the absence of the candidate compound. The candidate compound can then be identified as a modulator of ORFX expression based on this comparison. For example, when expression of ORFX mRNA or protein is greater (statistically significantly greater) in the presence of the candidate compound than in its absence, the candidate compound is identified as a stimulator of ORFX mRNA or protein expression. Alternatively, when expression of ORFX mRNA or protein is less (statistically significantly less) in the presence of the candidate

compound than in its absence, the candidate compound is identified as an inhibitor of ORFX mRNA or protein expression. The level of ORFX mRNA or protein expression in the cells can be determined by methods described herein for detecting ORFX mRNA or protein.

In yet another aspect of the invention, the ORFX proteins can be used as "bait proteins" in a two-hybrid assay or three hybrid assay (see, *e.g.*, U.S. Pat. No. 5,283,317; Zervos *et al.* (1993) Cell 72:223-232; Madura *et al.* (1993) J Biol Chem 268:12046-12054; Bartel *et al.* (1993) Biotechniques 14:920-924; Iwabuchi *et al.* (1993) Oncogene 8:1693-1696; and Brent WO94/10300), to identify other proteins that bind to or interact with ORFX ("ORFX-binding proteins" or "ORFX-bp") and modulate ORFX activity. Such ORFX-binding proteins are also likely to be involved in the propagation of signals by the ORFX proteins as, for example, upstream or downstream elements of the ORFX pathway.

The two-hybrid system is based on the modular nature of most transcription factors, which consist of separable DNA-binding and activation domains. Briefly, the assay utilizes two different DNA constructs. In one construct, the gene that codes for ORFX is fused to a gene encoding the DNA binding domain of a known transcription factor (*e.g.*, GAL-4). In the other construct, a DNA sequence, from a library of DNA sequences, that encodes an unidentified protein ("prey" or "sample") is fused to a gene that codes for the activation domain of the known transcription factor. If the "bait" and the "prey" proteins are able to interact, *in vivo*, forming a ORFX-dependent complex, the DNA-binding and activation domains of the transcription factor are brought into close proximity. This proximity allows transcription of a reporter gene (*e.g.*, LacZ) that is operably linked to a transcriptional regulatory site responsive to the transcription factor. Expression of the reporter gene can be detected and cell colonies containing the functional transcription factor can be isolated and used to obtain the cloned gene that encodes the protein which interacts with ORFX.

This invention further pertains to novel agents identified by the above-described screening assays and uses thereof for treatments as described herein.

Detection Assays

Portions or fragments of the cDNA sequences identified herein (and the corresponding complete gene sequences) can be used in numerous ways as polynucleotide reagents. For example, these sequences can be used to: (i) map their respective genes on a chromosome; and, thus, locate gene regions associated with genetic disease; (ii) identify an individual from a

minute biological sample (tissue typing); and (iii) aid in forensic identification of a biological sample.

The ORFX sequences of the present invention can also be used to identify individuals from minute biological samples. In this technique, an individual's genomic DNA is digested with one or more restriction enzymes, and probed on a Southern blot to yield unique bands for identification. The sequences of the present invention are useful as additional DNA markers for RFLP ("restriction fragment length polymorphisms," described in U.S. Pat. No. 5,272,057).

Furthermore, the sequences of the present invention can be used to provide an alternative technique that determines the actual base-by-base DNA sequence of selected portions of an individual's genome. Thus, the ORFX sequences described herein can be used to prepare two PCR primers from the 5' and 3' ends of the sequences. These primers can then be used to amplify an individual's DNA and subsequently sequence it.

Panels of corresponding DNA sequences from individuals, prepared in this manner, can provide unique individual identifications, as each individual will have a unique set of such DNA sequences due to allelic differences. The sequences of the present invention can be used to obtain such identification sequences from individuals and from tissue. The ORFX sequences of the invention uniquely represent portions of the human genome. Allelic variation occurs to some degree in the coding regions of these sequences, and to a greater degree in the noncoding regions. It is estimated that allelic variation between individual humans occurs with a frequency of about once per each 500 bases. Much of the allelic variation is due to single nucleotide polymorphisms (SNPs), which include restriction fragment length polymorphisms (RFLPs).

Each of the sequences described herein can, to some degree, be used as a standard against which DNA from an individual can be compared for identification purposes. Because greater numbers of polymorphisms occur in the noncoding regions, fewer sequences are necessary to differentiate individuals. The noncoding sequences of SEQ ID NO:2n-1 (wherein $n = 1$ to 3161), as described above, can comfortably provide positive individual identification with a panel of perhaps 10 to 1,000 primers that each yield a noncoding amplified sequence of 100 bases. If predicted coding sequences are used, a more appropriate number of primers for positive individual identification would be 500-2,000.

Predictive Medicine

The present invention also pertains to the field of predictive medicine in which diagnostic assays, prognostic assays, pharmacogenomics, and monitoring clinical trials are used for prognostic (predictive) purposes to thereby treat an individual prophylactically. Accordingly, one aspect of the present invention relates to diagnostic assays for determining ORFX protein and/or nucleic acid expression as well as ORFX activity, in the context of a biological sample (e.g., blood, serum, cells, tissue) to thereby determine whether an individual is afflicted with a disease or disorder, or is at risk of developing a disorder, associated with aberrant ORFX expression or activity. The invention also provides for prognostic (or predictive) assays for determining whether an individual is at risk of developing a disorder associated with ORFX protein, nucleic acid expression or activity. For example, mutations in a ORFX gene can be assayed in a biological sample. Such assays can be used for prognostic or predictive purpose to thereby prophylactically treat an individual prior to the onset of a disorder characterized by or associated with ORFX protein, nucleic acid expression or activity.

Another aspect of the invention provides methods for determining ORFX protein, nucleic acid expression or ORFX activity in an individual to thereby select appropriate therapeutic or prophylactic agents for that individual (referred to herein as "pharmacogenomics"). Pharmacogenomics allows for the selection of agents (e.g., drugs) for therapeutic or prophylactic treatment of an individual based on the genotype of the individual (e.g., the genotype of the individual examined to determine the ability of the individual to respond to a particular agent.)

Yet another aspect of the invention pertains to monitoring the influence of agents (e.g., drugs, compounds) on the expression or activity of ORFX in clinical trials.

Use of Partial ORFX Sequences in Forensic Biology

DNA-based identification techniques can also be used in forensic biology. Forensic biology is a scientific field employing genetic typing of biological evidence found at a crime scene as a means for positively identifying, for example, a perpetrator of a crime. To make such an identification, PCR technology can be used to amplify DNA sequences taken from very small biological samples such as tissues, e.g., hair or skin, or body fluids, e.g., blood, saliva, or semen

found at a crime scene. The amplified sequence can then be compared to a standard, thereby allowing identification of the origin of the biological sample.

The sequences of the present invention can be used to provide polynucleotide reagents, *e.g.*, PCR primers, targeted to specific loci in the human genome, that can enhance the reliability of DNA-based forensic identifications by, for example, providing another "identification marker" (*i.e.* another DNA sequence that is unique to a particular individual). As mentioned above, actual base sequence information can be used for identification as an accurate alternative to patterns formed by restriction enzyme generated fragments. Sequences targeted to noncoding regions of SEQ ID NOs:___ are particularly appropriate for this use as greater numbers of polymorphisms occur in the noncoding regions, making it easier to differentiate individuals using this technique. Examples of polynucleotide reagents include the ORFX sequences or portions thereof, *e.g.*, fragments derived from the noncoding regions of one or more of SEQ ID NO:2*n*-1 (where *n* = 1 to 3161), having a length of at least 20 bases, preferably at least 30 bases.

The ORFX sequences described herein can further be used to provide polynucleotide reagents, *e.g.*, labeled or label-able probes that can be used, for example, in an *in situ* hybridization technique, to identify a specific tissue, *e.g.*, brain tissue, etc. This can be very useful in cases where a forensic pathologist is presented with a tissue of unknown origin. Panels of such ORFX probes can be used to identify tissue by species and/or by organ type.

In a similar fashion, these reagents, *e.g.*, ORFX primers or probes can be used to screen tissue culture for contamination (*i.e.* screen for the presence of a mixture of different types of cells in a culture).

Predictive Medicine

The present invention also pertains to the field of predictive medicine in which diagnostic assays, prognostic assays, pharmacogenomics, and monitoring clinical trials are used for prognostic (predictive) purposes to thereby treat an individual prophylactically. Accordingly, one aspect of the present invention relates to diagnostic assays for determining ORFX protein and/or nucleic acid expression as well as ORFX activity, in the context of a biological sample (*e.g.*, blood, serum, cells, tissue) to thereby determine whether an individual is afflicted with a disease or disorder, or is at risk of developing a disorder, associated with aberrant ORFX expression or activity. The invention also provides for prognostic (or predictive) assays for determining whether an individual is at risk of developing a disorder associated with ORFX protein, nucleic

acid expression or activity. For example, mutations in a ORFX gene can be assayed in a biological sample. Such assays can be used for prognostic or predictive purpose to thereby prophylactically treat an individual prior to the onset of a disorder characterized by or associated with ORFX protein, nucleic acid expression or activity.

5 Another aspect of the invention provides methods for determining ORFX protein, nucleic acid expression or ORFX activity in an individual to thereby select appropriate therapeutic or prophylactic agents for that individual (referred to herein as "pharmacogenomics"). Pharmacogenomics allows for the selection of agents (e.g., drugs) for therapeutic or prophylactic treatment of an individual based on the genotype of the individual (e.g., the genotype of the individual examined to determine the ability of the individual to respond to a particular agent.)

10 Yet another aspect of the invention pertains to monitoring the influence of agents (e.g., drugs, compounds) on the expression or activity of ORFX in clinical trials.

These and other agents are described in further detail in the following sections.

Diagnostic Assays

15 Other conditions in which proliferation of cells plays a role include tumors, restenosis, psoriasis, Dupuytren's contracture, diabetic complications, Kaposi's sarcoma and rheumatoid arthritis.

20 An ORFX polypeptide may be used to identify an interacting polypeptide a sample or tissue. The method comprises contacting the sample or tissue with ORFX, allowing formation of a complex between the ORFX polypeptide and the interacting polypeptide, and detecting the complex, if present.

25 The proteins of the invention may be used to stimulate production of antibodies specifically binding the proteins. Such antibodies may be used in immunodiagnostic procedures to detect the occurrence of the protein in a sample. The proteins of the invention may be used to stimulate cell growth and cell proliferation in conditions in which such growth would be favorable. An example would be to counteract toxic side effects of chemotherapeutic agents on, for example, hematopoiesis and platelet formation, linings of the gastrointestinal tract, and hair follicles. They may also be used to stimulate new cell growth in neurological disorders including, for example, Alzheimer's disease. Alternatively, antagonistic treatments may be

30 administered in which an antibody specifically binding the ORFX-like proteins of the invention

would abrogate the specific growth-inducing effects of the proteins. Such antibodies may be useful, for example, in the treatment of proliferative disorders including various tumors and benign hyperplasias.

Polynucleotides or oligonucleotides corresponding to any one portion of the ORFX nucleic acids of SEQ ID NO:2*n*-1 (wherein *n* = 1 to 3161) may be used to detect DNA
5 containing a corresponding ORF gene, or detect the expression of a corresponding ORFX gene, or ORFX-like gene. For example, an ORFX nucleic acid expressed in a particular cell or tissue, as noted in Table 2, can be used to identify the presence of that particular cell type.

An exemplary method for detecting the presence or absence of ORFX in a biological
10 sample involves obtaining a biological sample from a test subject and contacting the biological sample with a compound or an agent capable of detecting ORFX protein or nucleic acid (e.g., mRNA, genomic DNA) that encodes ORFX protein such that the presence of ORFX is detected in the biological sample. An agent for detecting ORFX mRNA or genomic DNA is a labeled nucleic acid probe capable of hybridizing to ORFX mRNA or genomic DNA. The nucleic acid
15 probe can be, for example, a full-length ORFX nucleic acid, such as the nucleic acid of SEQ ID NO:2*n*-1 (wherein *n* = 1 to 3161), or a portion thereof, such as an oligonucleotide of at least 15, 30, 50, 100, 250 or 500 nucleotides in length and sufficient to specifically hybridize under stringent conditions to ORFX mRNA or genomic DNA, as described above. Other suitable probes for use in the diagnostic assays of the invention are described herein.

20 An agent for detecting ORFX protein is an antibody capable of binding to ORFX protein, preferably an antibody with a detectable label. Antibodies can be polyclonal, or more preferably, monoclonal. An intact antibody, or a fragment thereof (e.g., Fab or F(ab')₂) can be used. The term "labeled", with regard to the probe or antibody, is intended to encompass direct labeling of the probe or antibody by coupling (i.e., physically linking) a detectable substance to the probe or
25 antibody, as well as indirect labeling of the probe or antibody by reactivity with another reagent that is directly labeled. Examples of indirect labeling include detection of a primary antibody using a fluorescently labeled secondary antibody and end-labeling of a DNA probe with biotin such that it can be detected with fluorescently labeled streptavidin. The term "biological sample" is intended to include tissues, cells and biological fluids isolated from a subject, as well as
30 tissues, cells and fluids present within a subject. That is, the detection method of the invention can be used to detect ORFX mRNA, protein, or genomic DNA in a biological sample *in vitro* as

well as *in vivo*. For example, *in vitro* techniques for detection of ORFX mRNA include Northern hybridizations and *in situ* hybridizations. *In vitro* techniques for detection of ORFX protein include enzyme linked immunosorbent assays (ELISAs), Western blots, immunoprecipitations and immunofluorescence. *In vitro* techniques for detection of ORFX genomic DNA include Southern hybridizations. Furthermore, *in vivo* techniques for detection of ORFX protein include introducing into a subject a labeled anti-ORFX antibody. For example, the antibody can be labeled with a radioactive marker whose presence and location in a subject can be detected by standard imaging techniques.

In one embodiment, the biological sample contains protein molecules from the test subject. Alternatively, the biological sample can contain mRNA molecules from the test subject or genomic DNA molecules from the test subject. A preferred biological sample is a peripheral blood leukocyte sample isolated by conventional means from a subject.

In another embodiment, the methods further involve obtaining a control biological sample from a control subject, contacting the control sample with a compound or agent capable of detecting ORFX protein, mRNA, or genomic DNA, such that the presence of ORFX protein, mRNA or genomic DNA is detected in the biological sample, and comparing the presence of ORFX protein, mRNA or genomic DNA in the control sample with the presence of ORFX protein, mRNA or genomic DNA in the test sample.

The invention also encompasses kits for detecting the presence of ORFX in a biological sample. For example, the kit can comprise: a labeled compound or agent capable of detecting ORFX protein or mRNA in a biological sample; means for determining the amount of ORFX in the sample; and means for comparing the amount of ORFX in the sample with a standard. The compound or agent can be packaged in a suitable container. The kit can further comprise instructions for using the kit to detect ORFX protein or nucleic acid.

Prognostic Assays

The diagnostic methods described herein can furthermore be utilized to identify subjects having or at risk of developing a disease or disorder associated with aberrant ORFX expression or activity. For example, the assays described herein, such as the preceding diagnostic assays or the following assays, can be utilized to identify a subject having or at risk of developing a disorder associated with ORFX protein, nucleic acid expression or activity in, *e.g.*, proliferative or differentiative disorders such as hyperplasias, tumors, restenosis, psoriasis, Dupuytren's

contracture, diabetic complications, or rheumatoid arthritis, etc.; and glia-associated disorders such as cerebral lesions, diabetic neuropathies, cerebral edema, senile dementia, Alzheimer's disease, etc. Alternatively, the prognostic assays can be utilized to identify a subject having or at risk for developing a disease or disorder. Thus, the present invention provides a method for
5 identifying a disease or disorder associated with aberrant ORFX expression or activity in which a test sample is obtained from a subject and ORFX protein or nucleic acid (*e.g.*, mRNA, genomic DNA) is detected, wherein the presence of ORFX protein or nucleic acid is diagnostic for a subject having or at risk of developing a disease or disorder associated with aberrant ORFX expression or activity. As used herein, a "test sample" refers to a biological sample obtained
10 from a subject of interest. For example, a test sample can be a biological fluid (*e.g.*, serum), cell sample, or tissue.

Furthermore, the prognostic assays described herein can be used to determine whether a subject can be administered an agent (*e.g.*, an agonist, antagonist, peptidomimetic, protein, peptide, nucleic acid, small molecule, or other drug candidate) to treat a disease or disorder
15 associated with aberrant ORFX expression or activity. For example, such methods can be used to determine whether a subject can be effectively treated with an agent for a disorder, such as a proliferative disorder, differentiative disorder, glia-associated disorders, etc. Thus, the present invention provides methods for determining whether a subject can be effectively treated with an agent for a disorder associated with aberrant ORFX expression or activity in which a test sample
20 is obtained and ORFX protein or nucleic acid is detected (*e.g.*, wherein the presence of ORFX protein or nucleic acid is diagnostic for a subject that can be administered the agent to treat a disorder associated with aberrant ORFX expression or activity.)

The methods of the invention can also be used to detect genetic lesions in a ORFX gene, thereby determining if a subject with the lesioned gene is at risk for, or suffers from, a
25 proliferative disorder, differentiative disorder, glia-associated disorder, etc. In various embodiments, the methods include detecting, in a sample of cells from the subject, the presence or absence of a genetic lesion characterized by at least one of an alteration affecting the integrity of a gene encoding a ORFX-protein, or the mis-expression of the ORFX gene. For example, such genetic lesions can be detected by ascertaining the existence of at least one of (1) a deletion
30 of one or more nucleotides from a ORFX gene; (2) an addition of one or more nucleotides to a ORFX gene; (3) a substitution of one or more nucleotides of a ORFX gene, (4) a chromosomal

rearrangement of a ORFX gene; (5) an alteration in the level of a messenger RNA transcript of a ORFX gene, (6) aberrant modification of a ORFX gene, such as of the methylation pattern of the genomic DNA, (7) the presence of a non-wild type splicing pattern of a messenger RNA transcript of a ORFX gene, (8) a non-wild type level of a ORFX-protein, (9) allelic loss of a ORFX gene, and (10) inappropriate post-translational modification of a ORFX-protein. As described herein, there are a large number of assay techniques known in the art which can be used for detecting lesions in a ORFX gene. A preferred biological sample is a peripheral blood leukocyte sample isolated by conventional means from a subject. However, any biological sample containing nucleated cells may be used, including, for example, buccal mucosal cells.

10 In certain embodiments, detection of the lesion involves the use of a probe/primer in a polymerase chain reaction (PCR) (see, *e.g.*, U.S. Pat. Nos. 4,683,195 and 4,683,202), such as anchor PCR or RACE PCR, or, alternatively, in a ligation chain reaction (LCR) (see, *e.g.*, Landegran *et al.* (1988) *Science* 241:1077-1080; and Nakazawa *et al.* (1994) *PNAS* 91:360-364), the latter of which can be particularly useful for detecting point mutations in the ORFX-gene
15 (see Abravaya *et al.* (1995) *Nucl Acids Res* 23:675-682). This method can include the steps of collecting a sample of cells from a patient, isolating nucleic acid (*e.g.*, genomic, mRNA or both) from the cells of the sample, contacting the nucleic acid sample with one or more primers that specifically hybridize to a ORFX gene under conditions such that hybridization and amplification of the ORFX gene (if present) occurs, and detecting the presence or absence of an
20 amplification product, or detecting the size of the amplification product and comparing the length to a control sample. It is anticipated that PCR and/or LCR may be desirable to use as a preliminary amplification step in conjunction with any of the techniques used for detecting mutations described herein.

Alternative amplification methods include: self sustained sequence replication (Guatelli
25 *et al.*, 1990, *Proc Natl Acad Sci USA* 87:1874-1878), transcriptional amplification system (Kwoh, *et al.*, 1989, *Proc Natl Acad Sci USA* 86:1173-1177), Q-Beta Replicase (Lizardi *et al.*, 1988, *BioTechnology* 6:1197), or any other nucleic acid amplification method, followed by the detection of the amplified molecules using techniques well known to those of skill in the art. These detection schemes are especially useful for the detection of nucleic acid molecules if such
30 molecules are present in very low numbers.

In an alternative embodiment, mutations in a ORFX gene from a sample cell can be identified by alterations in restriction enzyme cleavage patterns. For example, sample and control DNA is isolated, amplified (optionally), digested with one or more restriction endonucleases, and fragment length sizes are determined by gel electrophoresis and compared.

5 Differences in fragment length sizes between sample and control DNA indicates mutations in the sample DNA. Moreover, the use of sequence specific ribozymes (see, for example, U.S. Pat. No. 5,493,531) can be used to score for the presence of specific mutations by development or loss of a ribozyme cleavage site.

In other embodiments, genetic mutations in ORFX can be identified by hybridizing a
10 sample and control nucleic acids, *e.g.*, DNA or RNA, to high density arrays containing hundreds or thousands of oligonucleotide probes (Cronin *et al.* (1996) *Human Mutation* 7: 244-255; Kozal *et al.* (1996) *Nature Medicine* 2: 753-759). For example, genetic mutations in ORFX can be identified in two dimensional arrays containing light-generated DNA probes as described in Cronin *et al.* above. Briefly, a first hybridization array of probes can be used to scan through
15 long stretches of DNA in a sample and control to identify base changes between the sequences by making linear arrays of sequential overlapping probes. This step allows the identification of point mutations. This step is followed by a second hybridization array that allows the characterization of specific mutations by using smaller, specialized probe arrays complementary to all variants or mutations detected. Each mutation array is composed of parallel probe sets, one
20 complementary to the wild-type gene and the other complementary to the mutant gene.

In yet another embodiment, any of a variety of sequencing reactions known in the art can be used to directly sequence the ORFX gene and detect mutations by comparing the sequence of the sample ORFX with the corresponding wild-type (control) sequence. Examples of sequencing reactions include those based on techniques developed by Maxim and Gilbert (1977) *PNAS*
25 74:560 or Sanger (1977) *PNAS* 74:5463. It is also contemplated that any of a variety of automated sequencing procedures can be utilized when performing the diagnostic assays (Naeve *et al.*, (1995) *Biotechniques* 19:448), including sequencing by mass spectrometry (see, *e.g.*, PCT International Publ. No. WO 94/16101; Cohen *et al.* (1996) *Adv Chromatogr* 36:127-162; and Griffin *et al.* (1993) *Appl Biochem Biotechnol* 38:147-159).

30 Other methods for detecting mutations in the ORFX gene include methods in which protection from cleavage agents is used to detect mismatched bases in RNA/RNA or RNA/DNA

heteroduplexes (Myers *et al.* (1985) *Science* 230:1242). In general, the art technique of "mismatch cleavage" starts by providing heteroduplexes of formed by hybridizing (labeled) RNA or DNA containing the wild-type ORFX sequence with potentially mutant RNA or DNA obtained from a tissue sample. The double-stranded duplexes are treated with an agent that cleaves single-stranded regions of the duplex such as which will exist due to basepair mismatches between the control and sample strands. For instance, RNA/DNA duplexes can be treated with RNase and DNA/DNA hybrids treated with S1 nuclease to enzymatically digesting the mismatched regions. In other embodiments, either DNA/DNA or RNA/DNA duplexes can be treated with hydroxylamine or osmium tetroxide and with piperidine in order to digest mismatched regions. After digestion of the mismatched regions, the resulting material is then separated by size on denaturing polyacrylamide gels to determine the site of mutation. See, for example, Cotton *et al* (1988) *Proc Natl Acad Sci USA* 85:4397; Saleeba *et al* (1992) *Methods Enzymol* 217:286-295. In an embodiment, the control DNA or RNA can be labeled for detection.

In still another embodiment, the mismatch cleavage reaction employs one or more proteins that recognize mismatched base pairs in double-stranded DNA (so called "DNA mismatch repair" enzymes) in defined systems for detecting and mapping point mutations in ORFX cDNAs obtained from samples of cells. For example, the mutY enzyme of *E. coli* cleaves A at G/A mismatches and the thymidine DNA glycosylase from HeLa cells cleaves T at G/T mismatches (Hsu *et al.* (1994) *Carcinogenesis* 15:1657-1662). According to an exemplary embodiment, a probe based on a ORFX sequence, *e.g.*, a wild-type ORFX sequence, is hybridized to a cDNA or other DNA product from a test cell(s). The duplex is treated with a DNA mismatch repair enzyme, and the cleavage products, if any, can be detected from electrophoresis protocols or the like. See, for example, U.S. Pat. No. 5,459,039.

In other embodiments, alterations in electrophoretic mobility will be used to identify mutations in ORFX genes. For example, single strand conformation polymorphism (SSCP) may be used to detect differences in electrophoretic mobility between mutant and wild type nucleic acids (Orita *et al.* (1989) *Proc Natl Acad Sci USA*: 86:2766, see also Cotton (1993) *Mutat Res* 285:125-144; Hayashi (1992) *Genet Anal Tech Appl* 9:73-79). Single-stranded DNA fragments of sample and control ORFX nucleic acids will be denatured and allowed to renature. The secondary structure of single-stranded nucleic acids varies according to sequence, the resulting

alteration in electrophoretic mobility enables the detection of even a single base change. The DNA fragments may be labeled or detected with labeled probes. The sensitivity of the assay may be enhanced by using RNA, rather than DNA, in which the secondary structure is more sensitive to a change in sequence. In one embodiment, the subject method utilizes heteroduplex analysis to separate double stranded heteroduplex molecules on the basis of changes in electrophoretic mobility. See, *e.g.*, Keen *et al.* (1991) *Trends Genet* 7:5.

In yet another embodiment the movement of mutant or wild-type fragments in polyacrylamide gels containing a gradient of denaturant is assayed using denaturing gradient gel electrophoresis (DGGE). See, *e.g.*, Myers *et al.* (1985) *Nature* 313:495. When DGGE is used as the method of analysis, DNA will be modified to insure that it does not completely denature, for example by adding a GC clamp of approximately 40 bp of high-melting GC-rich DNA by PCR. In a further embodiment, a temperature gradient is used in place of a denaturing gradient to identify differences in the mobility of control and sample DNA. See, *e.g.*, Rosenbaum and Reissner (1987) *Biophys Chem* 265:12753.

Examples of other techniques for detecting point mutations include, but are not limited to, selective oligonucleotide hybridization, selective amplification, or selective primer extension. For example, oligonucleotide primers may be prepared in which the known mutation is placed centrally and then hybridized to target DNA under conditions that permit hybridization only if a perfect match is found. See, *e.g.*, Saiki *et al.* (1986) *Nature* 324:163; Saiki *et al.* (1989) *Proc Natl Acad. Sci USA* 86:6230. Such allele specific oligonucleotides are hybridized to PCR amplified target DNA or a number of different mutations when the oligonucleotides are attached to the hybridizing membrane and hybridized with labeled target DNA.

Alternatively, allele specific amplification technology that depends on selective PCR amplification may be used in conjunction with the instant invention. Oligonucleotides used as primers for specific amplification may carry the mutation of interest in the center of the molecule (so that amplification depends on differential hybridization) (Gibbs *et al.* (1989) *Nucleic Acids Res* 17:2437-2448) or at the extreme 3' end of one primer where, under appropriate conditions, mismatch can prevent, or reduce polymerase extension (Prossner (1993) *Tibtech* 11:238). In addition it may be desirable to introduce a novel restriction site in the region of the mutation to create cleavage-based detection. See, *e.g.*, Gasparini *et al.* (1992) *Mol Cell Probes* 6:1. It is anticipated that in certain embodiments amplification may also be performed using Taq ligase

for amplification. See, *e.g.*, Barany (1991) *Proc Natl Acad Sci USA* 88:189. In such cases, ligation will occur only if there is a perfect match at the 3' end of the 5' sequence, making it possible to detect the presence of a known mutation at a specific site by looking for the presence or absence of amplification.

5 The methods described herein may be performed, for example, by utilizing pre-packaged diagnostic kits comprising at least one probe nucleic acid or antibody reagent described herein, which may be conveniently used, *e.g.*, in clinical settings to diagnose patients exhibiting symptoms or family history of a disease or illness involving a ORFX gene.

10 Furthermore, any cell type or tissue, preferably peripheral blood leukocytes, in which ORFX is expressed may be utilized in the prognostic assays described herein. However, any biological sample containing nucleated cells may be used, including, for example, buccal mucosal cells.

Pharmacogenomics

15 Agents, or modulators that have a stimulatory or inhibitory effect on ORFX activity (*e.g.*, ORFX gene expression), as identified by a screening assay described herein can be administered to individuals to treat (prophylactically or therapeutically) disorders (*e.g.*, neurological, cancer-related or gestational disorders) associated with aberrant ORFX activity. In conjunction with such treatment, the pharmacogenomics (*i.e.*, the study of the relationship between an individual's genotype and that individual's response to a foreign compound or drug) of the individual may be
20 considered. Differences in metabolism of therapeutics can lead to severe toxicity or therapeutic failure by altering the relation between dose and blood concentration of the pharmacologically active drug. Thus, the pharmacogenomics of the individual permits the selection of effective agents (*e.g.*, drugs) for prophylactic or therapeutic treatments based on a consideration of the individual's genotype. Such pharmacogenomics can further be used to determine appropriate
25 dosages and therapeutic regimens. Accordingly, the activity of ORFX protein, expression of ORFX nucleic acid, or mutation content of ORFX genes in an individual can be determined to thereby select appropriate agent(s) for therapeutic or prophylactic treatment of the individual.

30 Pharmacogenomics deals with clinically significant hereditary variations in the response to drugs due to altered drug disposition and abnormal action in affected persons. See *e.g.*, Eichelbaum, 1996, *Clin Exp Pharmacol Physiol*, 23:983-985 and Linder, 1997, *Clin Chem*, 43:254-266. In general, two types of pharmacogenetic conditions can be differentiated. Genetic

conditions transmitted as a single factor altering the way drugs act on the body (altered drug action) or genetic conditions transmitted as single factors altering the way the body acts on drugs (altered drug metabolism). These pharmacogenetic conditions can occur either as rare defects or as polymorphisms. For example, glucose-6-phosphate dehydrogenase (G6PD) deficiency is a
5 common inherited enzymopathy in which the main clinical complication is haemolysis after ingestion of oxidant drugs (anti-malarials, sulfonamides, analgesics, nitrofurans) and consumption of fava beans.

As an illustrative embodiment, the activity of drug metabolizing enzymes is a major determinant of both the intensity and duration of drug action. The discovery of genetic
10 polymorphisms of drug metabolizing enzymes (*e.g.*, N-acetyltransferase 2 (NAT 2) and cytochrome P450 enzymes CYP2D6 and CYP2C19) has provided an explanation as to why some patients do not obtain the expected drug effects or show exaggerated drug response and serious toxicity after taking the standard and safe dose of a drug. These polymorphisms are expressed in two phenotypes in the population, the extensive metabolizer (EM) and poor
15 metabolizer (PM). The prevalence of PM is different among different populations. For example, the gene coding for CYP2D6 is highly polymorphic and several mutations have been identified in PM, which all lead to the absence of functional CYP2D6. Poor metabolizers of CYP2D6 and CYP2C19 quite frequently experience exaggerated drug response and side effects when they receive standard doses. If a metabolite is the active therapeutic moiety, PM show no therapeutic
20 response, as demonstrated for the analgesic effect of codeine mediated by its CYP2D6-formed metabolite morphine. The other extreme are the so called ultra-rapid metabolizers who do not respond to standard doses. Recently, the molecular basis of ultra-rapid metabolism has been identified to be due to CYP2D6 gene amplification.

Thus, the activity of ORFX protein, expression of ORFX nucleic acid, or mutation
25 content of ORFX genes in an individual can be determined to thereby select appropriate agent(s) for therapeutic or prophylactic treatment of the individual. In addition, pharmacogenetic studies can be used to apply genotyping of polymorphic alleles encoding drug-metabolizing enzymes to the identification of an individual's drug responsiveness phenotype. This knowledge, when applied to dosing or drug selection, can avoid adverse reactions or therapeutic failure and thus
30 enhance therapeutic or prophylactic efficiency when treating a subject with a ORFX modulator, such as a modulator identified by one of the exemplary screening assays described herein.

Monitoring Clinical Efficacy

Monitoring the influence of agents (*e.g.*, drugs, compounds) on the expression or activity of ORFX (*e.g.*, the ability to modulate aberrant cell proliferation and/or differentiation) can be applied in basic drug screening and in clinical trials. For example, the effectiveness of an agent
5 determined by a screening assay as described herein to increase ORFX gene expression, protein levels, or upregulate ORFX activity, can be monitored in clinical trials of subjects exhibiting decreased ORFX gene expression, protein levels, or downregulated ORFX activity.

Alternatively, the effectiveness of an agent determined by a screening assay to decrease ORFX gene expression, protein levels, or downregulate ORFX activity, can be monitored in clinical
10 trials of subjects exhibiting increased ORFX gene expression, protein levels, or upregulated ORFX activity. In such clinical trials, the expression or activity of ORFX and, preferably, other genes that have been implicated in, for example, a proliferative or neurological disorder, can be used as a "read out" or marker of the responsiveness of a particular cell.

For example, genes, including ORFX, that are modulated in cells by treatment with an
15 agent (*e.g.*, compound, drug or small molecule) that modulates ORFX activity (*e.g.*, identified in a screening assay as described herein) can be identified. Thus, to study the effect of agents on cellular proliferation disorders, for example, in a clinical trial, cells can be isolated and RNA prepared and analyzed for the levels of expression of ORFX and other genes implicated in the disorder. The levels of gene expression (*i.e.*, a gene expression pattern) can be quantified by
20 Northern blot analysis or RT-PCR, as described herein, or alternatively by measuring the amount of protein produced, by one of the methods as described herein, or by measuring the levels of activity of ORFX or other genes. In this way, the gene expression pattern can serve as a marker, indicative of the physiological response of the cells to the agent. Accordingly, this response state may be determined before, and at various points during, treatment of the individual with the
25 agent.

In one embodiment, the invention provides a method for monitoring the effectiveness of treatment of a subject with an agent (*e.g.*, an agonist, antagonist, protein, peptide, nucleic acid, peptidomimetic, small molecule, or other drug candidate identified by the screening assays described herein) comprising the steps of (i) obtaining a pre-administration sample from a
30 subject prior to administration of the agent; (ii) detecting the level of expression of a ORFX protein, mRNA, or genomic DNA in the preadministration sample; (iii) obtaining one or more

post-administration samples from the subject; (iv) detecting the level of expression or activity of the ORFX protein, mRNA, or genomic DNA in the post-administration samples; (v) comparing the level of expression or activity of the ORFX protein, mRNA, or genomic DNA in the pre-administration sample with the ORFX protein, mRNA, or genomic DNA in the post administration sample or samples; and (vi) altering the administration of the agent to the subject accordingly. For example, increased administration of the agent may be desirable to increase the expression or activity of ORFX to higher levels than detected, *i.e.*, to increase the effectiveness of the agent. Alternatively, decreased administration of the agent may be desirable to decrease expression or activity of ORFX to lower levels than detected, *i.e.*, to decrease the effectiveness of the agent.

Methods of Treatment

The present invention provides for both prophylactic and therapeutic methods of treating a subject at risk of (or susceptible to) a disorder or having a disorder associated with aberrant ORFX expression or activity.

Diseases and disorders that are characterized by increased (relative to a subject not suffering from the disease or disorder) levels or biological activity may be treated with Therapeutics that antagonize (*i.e.*, reduce or inhibit) activity. Therapeutics that antagonize activity may be administered in a therapeutic or prophylactic manner. Therapeutics that may be utilized include, but are not limited to, (i) a ORFX polypeptide, or analogs, derivatives, fragments or homologs thereof; (ii) antibodies to a ORFX peptide; (iii) nucleic acids encoding a ORFX peptide; (iv) administration of antisense nucleic acid and nucleic acids that are "dysfunctional" (*i.e.*, due to a heterologous insertion within the coding sequences of coding sequences to a ORFX peptide) that are utilized to "knockout" endogenous function of a ORFX peptide by homologous recombination (see, *e.g.*, Capecchi, 1989, *Science* 244: 1288-1292); or (v) modulators (*i.e.*, inhibitors, agonists and antagonists, including additional peptide mimetic of the invention or antibodies specific to a peptide of the invention) that alter the interaction between a ORFX peptide and its binding partner.

Diseases and disorders that are characterized by decreased (relative to a subject not suffering from the disease or disorder) levels or biological activity may be treated with Therapeutics that increase (*i.e.*, are agonists to) activity. Therapeutics that upregulate activity may be administered in a therapeutic or prophylactic manner. Therapeutics that may be utilized

include, but are not limited to, a ORFX peptide, or analogs, derivatives, fragments or homologs thereof; or an agonist that increases bioavailability.

Increased or decreased levels can be readily detected by quantifying peptide and/or RNA, by obtaining a patient tissue sample (*e.g.*, from biopsy tissue) and assaying it *in vitro* for RNA or peptide levels, structure and/or activity of the expressed peptides (or mRNAs of a ORFX peptide). Methods that are well-known within the art include, but are not limited to, immunoassays (*e.g.*, by Western blot analysis, immunoprecipitation followed by sodium dodecyl sulfate (SDS) polyacrylamide gel electrophoresis, immunocytochemistry, etc.) and/or hybridization assays to detect expression of mRNAs (*e.g.*, Northern assays, dot blots, *in situ* hybridization, etc.).

In one aspect, the invention provides a method for preventing, in a subject, a disease or condition associated with an aberrant ORFX expression or activity, by administering to the subject an agent that modulates ORFX expression or at least one ORFX activity. Subjects at risk for a disease that is caused or contributed to by aberrant ORFX expression or activity can be identified by, for example, any or a combination of diagnostic or prognostic assays as described herein. Administration of a prophylactic agent can occur prior to the manifestation of symptoms characteristic of the ORFX aberrancy, such that a disease or disorder is prevented or, alternatively, delayed in its progression. Depending on the type of ORFX aberrancy, for example, a ORFX agonist or ORFX antagonist agent can be used for treating the subject. The appropriate agent can be determined based on screening assays described herein.

Another aspect of the invention pertains to methods of modulating ORFX expression or activity for therapeutic purposes. The modulatory method of the invention involves contacting a cell with an agent that modulates one or more of the activities of ORFX protein activity associated with the cell. An agent that modulates ORFX protein activity can be an agent as described herein, such as a nucleic acid or a protein, a naturally-occurring cognate ligand of a ORFX protein, a peptide, a ORFX peptidomimetic, or other small molecule. In one embodiment, the agent stimulates one or more ORFX protein activity. Examples of such stimulatory agents include active ORFX protein and a nucleic acid molecule encoding ORFX that has been introduced into the cell. In another embodiment, the agent inhibits one or more ORFX protein activity. Examples of such inhibitory agents include antisense ORFX nucleic acid molecules and anti-ORFX antibodies. These modulatory methods can be performed *in vitro*

(e.g., by culturing the cell with the agent) or, alternatively, *in vivo* (e.g., by administering the agent to a subject). As such, the present invention provides methods of treating an individual afflicted with a disease or disorder characterized by aberrant expression or activity of a ORFX protein or nucleic acid molecule. In one embodiment, the method involves administering an agent (e.g., an agent identified by a screening assay described herein), or combination of agents that modulates (e.g., upregulates or downregulates) ORFX expression or activity. In another embodiment, the method involves administering a ORFX protein or nucleic acid molecule as therapy to compensate for reduced or aberrant ORFX expression or activity.

Determination of the Biological Effect of a Therapeutic

In various embodiments of the present invention, suitable *in vitro* or *in vivo* assays are utilized to determine the effect of a specific Therapeutic and whether its administration is indicated for treatment of the affected tissue.

In various specific embodiments, *in vitro* assays may be performed with representative cells of the type(s) involved in the patient's disorder, to determine if a given Therapeutic exerts the desired effect upon the cell type(s). Compounds for use in therapy may be tested in suitable animal model systems including, but not limited to rats, mice, chicken, cows, monkeys, rabbits, and the like, prior to testing in human subjects. Similarly, for *in vivo* testing, any of the animal model system known in the art may be used prior to administration to human subjects.

Malignancies

Some ORFX polypeptides are expressed in cancerous cells (see, e.g., Tables 1 and 2). Accordingly, the corresponding ORF protein is involved in the regulation of cell proliferation. Accordingly, Therapeutics of the present invention may be useful in the therapeutic or prophylactic treatment of diseases or disorders that are associated with cell hyperproliferation and/or loss of control of cell proliferation (e.g., cancers, malignancies and tumors). For a review of such hyperproliferation disorders, see e.g., Fishman, *et al.*, 1985. MEDICINE, 2nd ed., J.B. Lippincott Co., Philadelphia, PA.

Therapeutics of the present invention may be assayed by any method known within the art for efficacy in treating or preventing malignancies and related disorders. Such assays include,

but are not limited to, *in vitro* assays utilizing transformed cells or cells derived from the patient's tumor, as well as *in vivo* assays using animal models of cancer or malignancies. Potentially effective Therapeutics are those that, for example, inhibit the proliferation of tumor-derived or transformed cells in culture or cause a regression of tumors in animal models, in comparison to the controls.

In the practice of the present invention, once a malignancy or cancer has been shown to be amenable to treatment by modulating (*i.e.*, inhibiting, antagonizing or agonizing) activity, that cancer or malignancy may subsequently be treated or prevented by the administration of a Therapeutic that serves to modulate protein function.

10 Premalignant conditions

The Therapeutics of the present invention that are effective in the therapeutic or prophylactic treatment of cancer or malignancies may also be administered for the treatment of pre-malignant conditions and/or to prevent the progression of a pre-malignancy to a neoplastic or malignant state. Such prophylactic or therapeutic use is indicated in conditions known or suspected of preceding progression to neoplasia or cancer, in particular, where non-neoplastic cell growth consisting of hyperplasia, metaplasia or, most particularly, dysplasia has occurred. For a review of such abnormal cell growth see *e.g.*, Robbins & Angell, 1976. BASIC PATHOLOGY, 2nd ed., W.B. Saunders Co., Philadelphia, PA.

Hyperplasia is a form of controlled cell proliferation involving an increase in cell number in a tissue or organ, without significant alteration in its structure or function. For example, it has been demonstrated that endometrial hyperplasia often precedes endometrial cancer. Metaplasia is a form of controlled cell growth in which one type of mature or fully differentiated cell substitutes for another type of mature cell. Metaplasia may occur in epithelial or connective tissue cells. Dysplasia is generally considered a precursor of cancer, and is found mainly in the epithelia. Dysplasia is the most disorderly form of non-neoplastic cell growth, and involves a loss in individual cell uniformity and in the architectural orientation of cells. Dysplasia characteristically occurs where there exists chronic irritation or inflammation, and is often found in the cervix, respiratory passages, oral cavity, and gall bladder.

Alternatively, or in addition to the presence of abnormal cell growth characterized as hyperplasia, metaplasia, or dysplasia, the presence of one or more characteristics of a transformed or malignant phenotype displayed either *in vivo* or *in vitro* within a cell sample derived from a patient, is indicative of the desirability of prophylactic/therapeutic administration of a Therapeutic that possesses the ability to modulate activity of An aforementioned protein.

5 Characteristics of a transformed phenotype include, but are not limited to: (i) morphological changes; (ii) looser substratum attachment; (iii) loss of cell-to-cell contact inhibition; (iv) loss of anchorage dependence; (v) protease release; (vi) increased sugar transport; (vii) decreased serum requirement; (viii) expression of fetal antigens, (ix) disappearance of the 250 kDal cell-surface

10 protein, and the like. See *e.g.*, Richards, *et al.*, 1986. MOLECULAR PATHOLOGY, W.B. Saunders Co., Philadelphia, PA.

In a specific embodiment of the present invention, a patient that exhibits one or more of the following predisposing factors for malignancy is treated by administration of an effective amount of a Therapeutic: (i) a chromosomal translocation associated with a malignancy (*e.g.*, the

15 Philadelphia chromosome (*bcr/abl*) for chronic myelogenous leukemia and t(14;18) for follicular lymphoma, etc.); (ii) familial polyposis or Gardner's syndrome (possible forerunners of colon cancer); (iii) monoclonal gammopathy of undetermined significance (a possible precursor of multiple myeloma) and (iv) a first degree kinship with persons having a cancer or pre-cancerous disease showing a Mendelian (genetic) inheritance pattern (*e.g.*, familial polyposis of the colon,

20 Gardner's syndrome, hereditary exostosis, polyendocrine adenomatosis, Peutz-Jeghers syndrome, neurofibromatosis of Von Recklinghausen, medullary thyroid carcinoma with amyloid production and pheochromocytoma, retinoblastoma, carotid body tumor, cutaneous melanocarcinoma, intraocular melanocarcinoma, xeroderma pigmentosum, ataxia telangiectasia, Chediak-Higashi syndrome, albinism, Fanconi's aplastic anemia and Bloom's syndrome).

25 In another embodiment, a Therapeutic of the present invention is administered to a human patient to prevent the progression to breast, colon, lung, pancreatic, or uterine cancer, or melanoma or sarcoma.

Hyperproliferative and dysproliferative disorders

In one embodiment of the present invention, a Therapeutic is administered in the therapeutic or prophylactic treatment of hyperproliferative or benign dysproliferative disorders. The efficacy in treating or preventing hyperproliferative diseases or disorders of a Therapeutic of
5 the present invention may be assayed by any method known within the art. Such assays include *in vitro* cell proliferation assays, *in vitro* or *in vivo* assays using animal models of hyperproliferative diseases or disorders, or the like. Potentially effective Therapeutics may, for example, promote cell proliferation in culture or cause growth or cell proliferation in animal models in comparison to controls.

10 Specific embodiments of the present invention are directed to the treatment or prevention of cirrhosis of the liver (a condition in which scarring has overtaken normal liver regeneration processes); treatment of keloid (hypertrophic scar) formation causing disfiguring of the skin in which the scarring process interferes with normal renewal; psoriasis (a common skin condition characterized by excessive proliferation of the skin and delay in proper cell fate determination);
15 benign tumors; fibrocystic conditions and tissue hypertrophy (*e.g.*, benign prostatic hypertrophy).

Neurodegenerative disorders

Some ORFX proteins are found in cell types have been implicated in the deregulation of cellular maturation and apoptosis, which are both characteristic of neurodegenerative disease. Accordingly, Therapeutics of the invention, particularly but not limited to those that modulate
20 (or supply) activity of an aforementioned protein, may be effective in treating or preventing neurodegenerative disease. Therapeutics of the present invention that modulate the activity of an aforementioned protein involved in neurodegenerative disorders can be assayed by any method known in the art for efficacy in treating or preventing such neurodegenerative diseases and disorders. Such assays include *in vitro* assays for regulated cell maturation or inhibition of
25 apoptosis or *in vivo* assays using animal models of neurodegenerative diseases or disorders, or any of the assays described below. Potentially effective Therapeutics, for example but not by way of limitation, promote regulated cell maturation and prevent cell apoptosis in culture, or reduce neurodegeneration in animal models in comparison to controls.

Once a neurodegenerative disease or disorder has been shown to be amenable to treatment by modulation activity, that neurodegenerative disease or disorder can be treated or prevented by administration of a Therapeutic that modulates activity. Such diseases include all degenerative disorders involved with aging, especially osteoarthritis and neurodegenerative disorders.

Disorders related to organ transplantation

Some ORFX can be associated with disorders related to organ transplantation, in particular but not limited to organ rejection. Therapeutics of the invention, particularly those that modulate (or supply) activity, may be effective in treating or preventing diseases or disorders related to organ transplantation. Therapeutics of the invention (particularly Therapeutics that modulate the levels or activity of an aforementioned protein) can be assayed by any method known in the art for efficacy in treating or preventing such diseases and disorders related to organ transplantation. Such assays include *in vitro* assays for using cell culture models as described below, or *in vivo* assays using animal models of diseases and disorders related to organ transplantation, see *e.g.*, below. Potentially effective Therapeutics, for example but not by way of limitation, reduce immune rejection responses in animal models in comparison to controls.

Accordingly, once diseases and disorders related to organ transplantation are shown to be amenable to treatment by modulation of activity, such diseases or disorders can be treated or prevented by administration of a Therapeutic that modulates activity.

Cardiovascular Disease

GENX has been implicated in cardiovascular disorders, including in atherosclerotic plaque formation. Diseases such as cardiovascular disease, including cerebral thrombosis or hemorrhage, ischemic heart or renal disease, peripheral vascular disease, or thrombosis of other major vessel, and other diseases, including diabetes mellitus, hypertension, hypothyroidism, cholesterol ester storage disease, systemic lupus erythematosus, homocysteinemia, and familial protein or lipid processing diseases, and the like, are either directly or indirectly associated with atherosclerosis. Accordingly, Therapeutics of the invention, particularly those that modulate (or supply) activity or formation may be effective in treating or preventing

atherosclerosis-associated diseases or disorders. Therapeutics of the invention (particularly Therapeutics that modulate the levels or activity) can be assayed by any method known in the art, including those described below, for efficacy in treating or preventing such diseases and disorders.

- 5 A vast array of animal and cell culture models exist for processes involved in atherosclerosis. A limited and non-exclusive list of animal models includes knockout mice for premature atherosclerosis (Kurabayashi and Yazaki, 1996, *Int. Angiol.* 15: 187-194), transgenic mouse models of atherosclerosis (Kappel *et al.*, 1994, *FASEB J.* 8: 583-592), antisense oligonucleotide treatment of animal models (Callow, 1995, *Curr. Opin. Cardiol.* 10: 569-576),
10 transgenic rabbit models for atherosclerosis (Taylor, 1997, *Ann. N.Y. Acad. Sci.* 811: 146-152), hypercholesterolemic animal models (Rosenfeld, 1996, *Diabetes Res. Clin. Pract.* 30 Suppl.: 1-11), hyperlipidemic mice (Paigen *et al.*, 1994, *Curr. Opin. Lipidol.* 5: 258-264), and inhibition of lipoxygenase in animals (Sigal *et al.*, 1994, *Ann. N.Y. Acad. Sci.* 714: 211-224). In addition, *in vitro* cell models include but are not limited to monocytes exposed to low density lipoprotein
15 (Frostegard *et al.*, 1996, *Atherosclerosis* 121: 93-103), cloned vascular smooth muscle cells (Suttles *et al.*, 1995, *Exp. Cell Res.* 218: 331-338), endothelial cell-derived chemoattractant exposed T cells (Katz *et al.*, 1994, *J. Leukoc. Biol.* 55: 567-573), cultured human aortic endothelial cells (Farber *et al.*, 1992, *Am. J. Physiol.* 262: H1088-1085), and foam cell cultures (Libby *et al.*, 1996, *Curr Opin Lipidol* 7: 330-335). Potentially effective Therapeutics, for
20 example but not by way of limitation, reduce foam cell formation in cell culture models, or reduce atherosclerotic plaque formation in hypercholesterolemic mouse models of atherosclerosis in comparison to controls.

- Accordingly, once an atherosclerosis-associated disease or disorder has been shown to be amenable to treatment by modulation of activity or formation, that disease or disorder can be
25 treated or prevented by administration of a Therapeutic that modulates activity.

Cytokine and Cell Proliferation/Differentiation Activity

A GENX protein of the present invention may exhibit cytokine, cell proliferation (either inducing or inhibiting) or cell differentiation (either inducing or inhibiting) activity or may induce production of other cytokines in certain cell populations. Many protein factors discovered

to date, including all known cytokines, have exhibited activity in one or more factor dependent cell proliferation assays, and hence the assays serve as a convenient confirmation of cytokine activity. The activity of a protein of the present invention is evidenced by any one of a number of routine factor dependent cell proliferation assays for cell lines including, without limitation, 32D,
5 DA2, DA1G, T10, B9, B9/11, BaF3, MC9/G, M+ (preB M+), 2E8, RB5, DA1, 123, T1165, HT2, CTLL2, TF-1, Mo7e and CMK.

The activity of a protein of the invention may, among other means, be measured by the following methods: Assays for T-cell or thymocyte proliferation include without limitation those described in: CURRENT PROTOCOLS IN IMMUNOLOGY, Ed by Coligan *et al.*, Greene Publishing
10 Associates and Wiley-Interscience (Chapter 3 and Chapter 7); Takai *et al.*, *J Immunol* 137:3494-3500, 1986; Bertagnoli *et al.*, *J Immunol* 145:1706-1712, 1990; Bertagnoli *et al.*, *Cell Immunol* 133:327-341, 1991; Bertagnoli, *et al.*, *J Immunol* 149:3778-3783, 1992; Bowman *et al.*, *J Immunol* 152:1756-1761, 1994.

Assays for cytokine production and/or proliferation of spleen cells, lymph node cells or
15 thymocytes include, without limitation, those described by Kruisbeek and Shevach, In: CURRENT PROTOCOLS IN IMMUNOLOGY. Coligan *et al.*, eds. Vol 1, pp. 3.12.1-14, John Wiley and Sons, Toronto 1994; and by Schreiber, In: CURRENT PROTOCOLS IN IMMUNOLOGY. Coligan eds. Vol 1 pp. 6.8.1-8, John Wiley and Sons, Toronto 1994.

Assays for proliferation and differentiation of hematopoietic and lymphopoietic cells
20 include, without limitation, those described by Bottomly *et al.*, In: CURRENT PROTOCOLS IN IMMUNOLOGY. Coligan *et al.*, eds. Vol 1 pp. 6.3.1-6.3.12, John Wiley and Sons, Toronto 1991; deVries *et al.*, *J Exp Med* 173:1205-1211, 1991; Moreau *et al.*, *Nature* 336:690-692, 1988; Greenberger *et al.*, *Proc Natl Acad Sci U.S.A.* 80:2931-2938, 1983; Nordan, In: CURRENT PROTOCOLS IN IMMUNOLOGY. Coligan *et al.*, eds. Vol 1 pp. 6.6.1-5, John Wiley and Sons,
25 Toronto 1991; Smith *et al.*, *Proc Natl Acad Sci U.S.A.* 83:1857-1861, 1986; Measurement of human Interleukin 11-Bennett, *et al.* In: CURRENT PROTOCOLS IN IMMUNOLOGY. Coligan *et al.*, eds. Vol 1 pp. 6.15.1 John Wiley and Sons, Toronto 1991; Ciarletta, *et al.*, In: CURRENT PROTOCOLS IN IMMUNOLOGY. Coligan *et al.*, eds. Vol 1 pp. 6.13.1, John Wiley and Sons, Toronto 1991.

30 Assays for T-cell clone responses to antigens (which will identify, among others, proteins that affect APC-T cell interactions as well as direct T-cell effects by measuring proliferation and

cytokine production) include, without limitation, those described in: CURRENT PROTOCOLS IN IMMUNOLOGY. Coligan *et al.*, eds., Greene Publishing Associates and Wiley-Interscience (Chapter 3Chapter 6, Chapter 7); Weinberger *et al.*, *Proc Natl Acad Sci USA* 77:6091-6095, 1980; Weinberger *et al.*, *Eur J Immunol* 11:405-411, 1981; Takai *et al.*, *J Immunol* 137:3494-3500, 5 1986; Takai *et al.*, *J Immunol* 140:508-512, 1988.

Immune Stimulating or Suppressing Activity

A GENX protein of the present invention may also exhibit immune stimulating or immune suppressing activity, including without limitation the activities for which assays are described herein. A protein may be useful in the treatment of various immune deficiencies and disorders (including severe combined immunodeficiency (SCID)), *e.g.*, in regulating (up or 10 down) growth and proliferation of T and/or B lymphocytes, as well as effecting the cytolytic activity of NK cells and other cell populations. These immune deficiencies may be genetic or be caused by viral (*e.g.*, HIV) as well as bacterial or fungal infections, or may result from autoimmune disorders. More specifically, infectious diseases caused by viral, bacterial, fungal or 15 other infection may be treatable using a protein of the present invention, including infections by HIV, hepatitis viruses, herpesviruses, mycobacteria, *Leishmania* species., malaria species. and various fungal infections such as candidiasis. Of course, in this regard, a protein of the present invention may also be useful where a boost to the immune system generally may be desirable, *i.e.*, in the treatment of cancer.

20 Autoimmune disorders which may be treated using a protein of the present invention include, for example, connective tissue disease, multiple sclerosis, systemic lupus erythematosus, rheumatoid arthritis, autoimmune pulmonary inflammation, Guillain-Barre syndrome, autoimmune thyroiditis, insulin dependent diabetes mellitus, myasthenia gravis, graft-versus-host disease and autoimmune inflammatory eye disease. Such a protein of the present invention may 25 also to be useful in the treatment of allergic reactions and conditions, such as asthma (particularly allergic asthma) or other respiratory problems. Other conditions, in which immune suppression is desired (including, for example, organ transplantation), may also be treatable using a protein of the present invention.

30 Using the proteins of the invention it may also be possible to immune responses, in a number of ways. Down regulation may be in the form of inhibiting or blocking an immune response already in progress or may involve preventing the induction of an immune response.

The functions of activated T cells may be inhibited by suppressing T cell responses or by inducing specific tolerance in T cells, or both. Immunosuppression of T cell responses is generally an active, non-antigen-specific, process which requires continuous exposure of the T cells to the suppressive agent. Tolerance, which involves inducing non-responsiveness or energy in T cells, is distinguishable from immunosuppression in that it is generally antigen-specific and persists after exposure to the tolerizing agent has ceased. Operationally, tolerance can be demonstrated by the lack of a T cell response upon re-exposure to specific antigen in the absence of the tolerizing agent.

Down regulating or preventing one or more antigen functions (including without limitation B lymphocyte antigen functions (such as, for example, B7), *e.g.*, preventing high level lymphokine synthesis by activated T cells, will be useful in situations of tissue, skin and organ transplantation and in graft-versus-host disease (GVHD). For example, blockage of T cell function should result in reduced tissue destruction in tissue transplantation. Typically, in tissue transplants, rejection of the transplant is initiated through its recognition as foreign by T cells, followed by an immune reaction that destroys the transplant. The administration of a molecule which inhibits or blocks interaction of a B7 lymphocyte antigen with its natural ligand(s) on immune cells (such as a soluble, monomeric form of a peptide having B7-2 activity alone or in conjunction with a monomeric form of a peptide having an activity of another B lymphocyte antigen (*e.g.*, B7-1, B7-3) or blocking antibody), prior to transplantation can lead to the binding of the molecule to the natural ligand(s) on the immune cells without transmitting the corresponding costimulatory signal. Blocking B lymphocyte antigen function in this matter prevents cytokine synthesis by immune cells, such as T cells, and thus acts as an immunosuppressant. Moreover, the lack of costimulation may also be sufficient to energize the T cells, thereby inducing tolerance in a subject. Induction of long-term tolerance by B lymphocyte antigen-blocking reagents may avoid the necessity of repeated administration of these blocking reagents. To achieve sufficient immunosuppression or tolerance in a subject, it may also be necessary to block the function of B lymphocyte antigens.

The efficacy of particular blocking reagents in preventing organ transplant rejection or GVHD can be assessed using animal models that are predictive of efficacy in humans. Examples of appropriate systems which can be used include allogeneic cardiac grafts in rats and xenogeneic pancreatic islet cell grafts in mice, both of which have been used to examine the

immunosuppressive effects of CTLA4Ig fusion proteins *in vivo* as described in Lenschow *et al.*, Science 257:789-792 (1992) and Turka *et al.*, Proc Natl Acad Sci USA, 89:11102-11105 (1992). In addition, murine models of GVHD (see Paul ed., FUNDAMENTAL IMMUNOLOGY, Raven Press, New York, 1989, pp. 846-847) can be used to determine the effect of blocking B lymphocyte antigen function *in vivo* on the development of that disease.

Blocking antigen function may also be therapeutically useful for treating autoimmune diseases. Many autoimmune disorders are the result of inappropriate activation of T cells that are reactive against self tissue and which promote the production of cytokines and auto-antibodies involved in the pathology of the diseases. Preventing the activation of autoreactive T cells may reduce or eliminate disease symptoms. Administration of reagents which block costimulation of T cells by disrupting receptor:ligand interactions of B lymphocyte antigens can be used to inhibit T cell activation and prevent production of auto-antibodies or T cell-derived cytokines which may be involved in the disease process. Additionally, blocking reagents may induce antigen-specific tolerance of autoreactive T cells which could lead to long-term relief from the disease. The efficacy of blocking reagents in preventing or alleviating autoimmune disorders can be determined using a number of well-characterized animal models of human autoimmune diseases. Examples include murine experimental autoimmune encephalitis, systemic lupus erythematosus in MRL/lpr/lpr mice or NZB hybrid mice, murine autoimmune collagen arthritis, diabetes mellitus in NOD mice and BB rats, and murine experimental myasthenia gravis (see Paul ed., FUNDAMENTAL IMMUNOLOGY, Raven Press, New York, 1989, pp. 840-856).

Upregulation of an antigen function (preferably a B lymphocyte antigen function), as a means of up regulating immune responses, may also be useful in therapy. Upregulation of immune responses may be in the form of enhancing an existing immune response or eliciting an initial immune response. For example, enhancing an immune response through stimulating B lymphocyte antigen function may be useful in cases of viral infection. In addition, systemic viral diseases such as influenza, the common cold, and encephalitis might be alleviated by the administration of stimulatory forms of B lymphocyte antigens systemically.

Alternatively, anti-viral immune responses may be enhanced in an infected patient by removing T cells from the patient, costimulating the T cells *in vitro* with viral antigen-pulsed APCs either expressing a peptide of the present invention or together with a stimulatory form of a soluble peptide of the present invention and reintroducing the *in vitro* activated T cells into the

patient. Another method of enhancing anti-vital immune responses would be to isolate infected cells from a patient, transfect them with a nucleic acid encoding a protein of the present invention as described herein such that the cells express all or a portion of the protein on their surface, and reintroduce the transfected cells into the patient. The infected cells would now be capable of delivering a costimulatory signal to, and thereby activate, T cells *in vivo*.

5 In another application, up regulation or enhancement of antigen function (preferably B lymphocyte antigen function) may be useful in the induction of tumor immunity. Tumor cells (e.g., sarcoma, melanoma, lymphoma, leukemia, neuroblastoma, carcinoma) transfected with a nucleic acid encoding at least one peptide of the present invention can be administered to a
10 subject to overcome tumor-specific tolerance in the subject. If desired, the tumor cell can be transfected to express a combination of peptides. For example, tumor cells obtained from a patient can be transfected *ex vivo* with an expression vector directing the expression of a peptide having B7-2-like activity alone, or in conjunction with a peptide having B7-1-like activity and/or B7-3-like activity. The transfected tumor cells are returned to the patient to result in expression
15 of the peptides on the surface of the transfected cell. Alternatively, gene therapy techniques can be used to target a tumor cell for transfection *in vivo*.

The presence of the peptide of the present invention having the activity of a B lymphocyte antigen(s) on the surface of the tumor cell provides the necessary costimulation signal to T cells to induce a T cell mediated immune response against the transfected tumor cells.
20 In addition, tumor cells which lack MHC class I or MHC class II molecules, or which fail to reexpress sufficient amounts of MHC class I or MHC class II molecules, can be transfected with nucleic acid encoding all or a portion of (e.g., a cytoplasmic-domain truncated portion) of an MHC class I α chain protein and β_2 microglobulin protein or an MHC class II α chain protein and an MHC class II β chain protein to thereby express MHC class I or MHC class II proteins on
25 the cell surface. Expression of the appropriate class I or class II MHC in conjunction with a peptide having the activity of a B lymphocyte antigen (e.g., B7-1, B7-2, B7-3) induces a T cell mediated immune response against the transfected tumor cell. Optionally, a gene encoding an antisense construct which blocks expression of an MHC class II associated protein, such as the invariant chain, can also be cotransfected with a DNA encoding a peptide having the activity of a
30 B lymphocyte antigen to promote presentation of tumor associated antigens and induce tumor

specific immunity. Thus, the induction of a T cell mediated immune response in a human subject may be sufficient to overcome tumor-specific tolerance in the subject.

The activity of a protein of the invention may, among other means, be measured by the following methods: Suitable assays for thymocyte or splenocyte cytotoxicity include, without
5 limitation, those described In: CURRENT PROTOCOLS IN IMMUNOLOGY. Coligan *et al.*, eds. Greene Publishing Associates and Wiley-Interscience (Chapter 3, Chapter 7); Herrmann *et al.*, *Proc Natl Acad Sci USA* 78:2488-2492, 1981; Herrmann *et al.*, *J Immunol* 128:1968-1974, 1982; Handa *et al.*, *J Immunol* 135:1564-1572, 1985; Takai *et al.*, *J Immunol* 137:3494-3500, 1986; Takai *et al.*, *J Immunol* 140:508-512, 1988; Herrmann *et al.*, *Proc Natl Acad Sci USA* 78:2488-2492, 1981;
10 Herrmann *et al.*, *J Immunol* 128:1968-1974, 1982; Handa *et al.*, *J Immunol* 135:1564-1572, 1985; Takai *et al.*, *J Immunol* 137:3494-3500, 1986; Bowman *et al.*, *J Virology* 61:1992-1998; Takai *et al.*, *J Immunol* 140:508-512, 1988; Bertagnolli *et al.*, *Cell Immunol* 133:327-341, 1991; Brown *et al.*, *J Immunol* 153:3079-3092, 1994.

Assays for T-cell-dependent immunoglobulin responses and isotype switching (which
15 will identify, among others, proteins that modulate T-cell dependent antibody responses and that affect Th1/Th2 profiles) include, without limitation, those described in: Maliszewski, *J Immunol* 144:3028-3033, 1990; and Mond and Brunswick In: CURRENT PROTOCOLS IN IMMUNOLOGY. Coligan *et al.*, (eds.) Vol 1 pp. 3.8.1-3.8.16, John Wiley and Sons, Toronto 1994.

Mixed lymphocyte reaction (MLR) assays (which will identify, among others, proteins
20 that generate predominantly Th1 and CTL responses) include, without limitation, those described In: CURRENT PROTOCOLS IN IMMUNOLOGY. Coligan *et al.*, eds. Greene Publishing Associates and Wiley-Interscience (Chapter 3, Chapter 7); Takai *et al.*, *J Immunol* 137:3494-3500, 1986; Takai *et al.*, *J Immunol* 140:508-512, 1988; Bertagnolli *et al.*, *J Immunol* 149:3778-3783, 1992.

Dendritic cell-dependent assays (which will identify, among others, proteins expressed by
25 dendritic cells that activate naive T-cells) include, without limitation, those described in: Guery *et al.*, *J Immunol* 134:536-544, 1995; Inaba *et al.*, *J Exp Med* 173:549-559, 1991; Macatonia *et al.*, *J Immunol* 154:5071-5079, 1995; Porgador *et al.*, *J Exp Med* 182:255-260, 1995; Nair *et al.*, *J Virol* 67:4062-4069, 1993; Huang *et al.*, *Science* 264:961-965, 1994; Macatonia *et al.*, *J Exp Med* 169:1255-1264, 1989; Bhardwaj *et al.*, *J Clin Investig* 94:797-807, 1994; and Inaba *et al.*, *J*
30 *Exp Med* 172:631-640, 1990.

Assays for lymphocyte survival/apoptosis (which will identify, among others, proteins that prevent apoptosis after superantigen induction and proteins that regulate lymphocyte homeostasis) include, without limitation, those described in: Darzynkiewicz *et al.*, *Cytometry* 13:795-808, 1992; Gorczyca *et al.*, *Leukemia* 7:659-670, 1993; Gorczyca *et al.*, *Cancer Res* 53:1945-1951, 1993; Itoh *et al.*, *Cell* 66:233-243, 1991; Zacharchuk, *J Immunol* 145:4037-4045, 1990; Zamai *et al.*, *Cytometry* 14:891-897, 1993; Gorczyca *et al.*, *Internat J Oncol* 1:639-648, 1992.

Assays for proteins that influence early steps of T-cell commitment and development include, without limitation, those described in: Antica *et al.*, *Blood* 84:111-117, 1994; Fine *et al.*, *Cell Immunol* 155: 111-122, 1994; Galy *et al.*, *Blood* 85:2770-2778, 1995; Toki *et al.*, *Proc Nat Acad Sci USA* 88:7548-7551, 1991.

Hematopoiesis Regulating Activity

A GENX protein of the present invention may be useful in regulation of hematopoiesis and, consequently, in the treatment of myeloid or lymphoid cell deficiencies. Even marginal biological activity in support of colony forming cells or of factor-dependent cell lines indicates involvement in regulating hematopoiesis, *e.g.* in supporting the growth and proliferation of erythroid progenitor cells alone or in combination with other cytokines, thereby indicating utility, for example, in treating various anemias or for use in conjunction with irradiation/chemotherapy to stimulate the production of erythroid precursors and/or erythroid cells; in supporting the growth and proliferation of myeloid cells such as granulocytes and monocytes/macrophages (*i.e.*, traditional CSF activity) useful, for example, in conjunction with chemotherapy to prevent or treat consequent myelo-suppression; in supporting the growth and proliferation of megakaryocytes and consequently of platelets thereby allowing prevention or treatment of various platelet disorders such as thrombocytopenia, and generally for use in place of or complimentary to platelet transfusions; and/or in supporting the growth and proliferation of hematopoietic stem cells which are capable of maturing to any and all of the above-mentioned hematopoietic cells and therefore find therapeutic utility in various stem cell disorders (such as those usually treated with transplantation, including, without limitation, aplastic anemia and paroxysmal nocturnal hemoglobinuria), as well as in repopulating the stem cell compartment post irradiation/chemotherapy, either in-vivo or ex-vivo (*i.e.*, in conjunction with bone marrow

transplantation or with peripheral progenitor cell transplantation (homologous or heterologous)) as normal cells or genetically manipulated for gene therapy.

The activity of a protein of the invention may, among other means, be measured by the following methods:

5 Suitable assays for proliferation and differentiation of various hematopoietic lines are cited above.

 Assays for embryonic stem cell differentiation (which will identify, among others, proteins that influence embryonic differentiation hematopoiesis) include, without limitation, those described in: Johansson *et al.* *Cellular Biology* 15:141-151, 1995; Keller *et al.*, *Mol. Cell.*
10 *Biol.* 13:473-486, 1993; McClanahan *et al.*, *Blood* 81:2903-2915, 1993.

 Assays for stem cell survival and differentiation (which will identify, among others, proteins that regulate lympho-hematopoiesis) include, without limitation, those described in: Methylcellulose colony forming assays, Freshney, In: CULTURE OF HEMATOPOIETIC CELLS. Freshney, *et al.* (eds.) Vol pp. 265-268, Wiley-Liss, Inc., New York, N.Y. 1994; Hirayama *et al.*,
15 *Proc Natl Acad Sci USA* 89:5907-5911, 1992; McNiece and Briddeli, In: CULTURE OF HEMATOPOIETIC CELLS. Freshney, *et al.* (eds.) Vol pp. 23-39, Wiley-Liss, Inc., New York, N.Y. 1994; Neben *et al.*, *Exp Hematol* 22:353-359, 1994; Ploemacher, In: CULTURE OF HEMATOPOIETIC CELLS. Freshney, *et al.* eds. Vol pp. 1-21, Wiley-Liss, Inc., New York, N.Y. 1994; Spoonceret *al.*, In: CULTURE OF HEMATOPOIETIC CELLS. Freshney, *et al.*, (eds.) Vol pp.
20 163-179, Wiley-Liss, Inc., New York, N.Y. 1994; Sutherland, In: CULTURE OF HEMATOPOIETIC CELLS. Freshney, *et al.*, (eds.) Vol pp. 139-162, Wiley-Liss, Inc., New York, N.Y. 1994.

Tissue Growth Activity

 A GENX protein of the present invention also may have utility in compositions used for bone, cartilage, tendon, ligament and/or nerve tissue growth or regeneration, as well as for
25 wound healing and tissue repair and replacement, and in the treatment of burns, incisions and ulcers.

 A protein of the present invention, which induces cartilage and/or bone growth in circumstances where bone is not normally formed, has application in the healing of bone fractures and cartilage damage or defects in humans and other animals. Such a preparation
30 employing a protein of the invention may have prophylactic use in closed as well as open fracture reduction and also in the improved fixation of artificial joints. De novo bone formation

induced by an osteogenic agent contributes to the repair of congenital, trauma induced, or oncologic resection induced craniofacial defects, and also is useful in cosmetic plastic surgery.

A protein of this invention may also be used in the treatment of periodontal disease, and in other tooth repair processes. Such agents may provide an environment to attract bone-forming cells, stimulate growth of bone-forming cells or induce differentiation of progenitors of bone-forming cells. A protein of the invention may also be useful in the treatment of osteoporosis or osteoarthritis, such as through stimulation of bone and/or cartilage repair or by blocking inflammation or processes of tissue destruction (collagenase activity, osteoclast activity, etc.) mediated by inflammatory processes.

Another category of tissue regeneration activity that may be attributable to the protein of the present invention is tendon/ligament formation. A protein of the present invention, which induces tendon/ligament-like tissue or other tissue formation in circumstances where such tissue is not normally formed, has application in the healing of tendon or ligament tears, deformities and other tendon or ligament defects in humans and other animals. Such a preparation employing a tendon/ligament-like tissue inducing protein may have prophylactic use in preventing damage to tendon or ligament tissue, as well as use in the improved fixation of tendon or ligament to bone or other tissues, and in repairing defects to tendon or ligament tissue. De novo tendon/ligament-like tissue formation induced by a composition of the present invention contributes to the repair of congenital, trauma induced, or other tendon or ligament defects of other origin, and is also useful in cosmetic plastic surgery for attachment or repair of tendons or ligaments. The compositions of the present invention may provide an environment to attract tendon- or ligament-forming cells, stimulate growth of tendon- or ligament-forming cells, induce differentiation of progenitors of tendon- or ligament-forming cells, or induce growth of tendon/ligament cells or progenitors *ex vivo* for return *in vivo* to effect tissue repair. The compositions of the invention may also be useful in the treatment of tendonitis, carpal tunnel syndrome and other tendon or ligament defects. The compositions may also include an appropriate matrix and/or sequestering agent as a carrier as is well known in the art.

The protein of the present invention may also be useful for proliferation of neural cells and for regeneration of nerve and brain tissue, *i.e.* for the treatment of central and peripheral nervous system diseases and neuropathies, as well as mechanical and traumatic disorders, which involve degeneration, death or trauma to neural cells or nerve tissue. More specifically, a protein

may be used in the treatment of diseases of the peripheral nervous system, such as peripheral nerve injuries, peripheral neuropathy and localized neuropathies, and central nervous system diseases, such as Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic lateral sclerosis, and Shy-Drager syndrome. Further conditions which may be treated in accordance with the present invention include mechanical and traumatic disorders, such as spinal cord disorders, head trauma and cerebrovascular diseases such as stroke. Peripheral neuropathies resulting from chemotherapy or other medical therapies may also be treatable using a protein of the invention.

Proteins of the invention may also be useful to promote better or faster closure of non-healing wounds, including without limitation pressure ulcers, ulcers associated with vascular insufficiency, surgical and traumatic wounds, and the like.

It is expected that a protein of the present invention may also exhibit activity for generation or regeneration of other tissues, such as organs (including, for example, pancreas, liver, intestine, kidney, skin, endothelium), muscle (smooth, skeletal or cardiac) and vascular (including vascular endothelium) tissue, or for promoting the growth of cells comprising such tissues. Part of the desired effects may be by inhibition or modulation of fibrotic scarring to allow normal tissue to regenerate. A protein of the invention may also exhibit angiogenic activity.

A protein of the present invention may also be useful for gut protection or regeneration and treatment of lung or liver fibrosis, reperfusion injury in various tissues, and conditions resulting from systemic cytokine damage.

A protein of the present invention may also be useful for promoting or inhibiting differentiation of tissues described above from precursor tissues or cells; or for inhibiting the growth of tissues described above.

The activity of a protein of the invention may, among other means, be measured by the following methods:

Assays for tissue generation activity include, without limitation, those described in: International Patent Publication No. WO95/16035 (bone, cartilage, tendon); International Patent Publication No. WO95/05846 (nerve, neuronal); International Patent Publication No. WO91/07491 (skin, endothelium).

Assays for wound healing activity include, without limitation, those described in: Winter, EPIDERMAL WOUND HEALING, pp. 71-112 (Maibach and Rovee, eds.), Year Book Medical

Publishers, Inc., Chicago, as modified by Eaglstein and Menz, *J. Invest. Dermatol* 71:382-84 (1978).

Activin/Inhibin Activity

A GENX protein of the present invention may also exhibit activin- or inhibin-related activities. Inhibins are characterized by their ability to inhibit the release of follicle stimulating hormone (FSH), while activins are characterized by their ability to stimulate the release of follicle stimulating hormone (FSH). Thus, a protein of the present invention, alone or in heterodimers with a member of the inhibin family, may be useful as a contraceptive based on the ability of inhibins to decrease fertility in female mammals and decrease spermatogenesis in male mammals. Administration of sufficient amounts of other inhibins can induce infertility in these mammals. Alternatively, the protein of the invention, as a homodimer or as a heterodimer with other protein subunits of the inhibin-b group, may be useful as a fertility inducing therapeutic, based upon the ability of activin molecules in stimulating FSH release from cells of the anterior pituitary. See, for example, U.S. Pat. No. 4,798,885. A protein of the invention may also be useful for advancement of the onset of fertility in sexually immature mammals, so as to increase the lifetime reproductive performance of domestic animals such as cows, sheep and pigs.

The activity of a protein of the invention may, among other means, be measured by the following methods:

Assays for activin/inhibin activity include, without limitation, those described in: Vale *et al.*, *Endocrinology* 91:562-572, 1972; Ling *et al.*, *Nature* 321:779-782, 1986; Vale *et al.*, *Nature* 321:776-779, 1986; Mason *et al.*, *Nature* 318:659-663, 1985; Forage *et al.*, *Proc Natl Acad Sci USA* 83:3091-3095, 1986.

Chemotactic/Chemokinetic Activity

A protein of the present invention may have chemotactic or chemokinetic activity (*e.g.*, act as a chemokine) for mammalian cells, including, for example, monocytes, fibroblasts, neutrophils, T-cells, mast cells, eosinophils, epithelial and/or endothelial cells. Chemotactic and chemokinetic proteins can be used to mobilize or attract a desired cell population to a desired site of action. Chemotactic or chemokinetic proteins provide particular advantages in treatment of wounds and other trauma to tissues, as well as in treatment of localized infections. For example,

attraction of lymphocytes, monocytes or neutrophils to tumors or sites of infection may result in improved immune responses against the tumor or infecting agent.

A protein or peptide has chemotactic activity for a particular cell population if it can stimulate, directly or indirectly, the directed orientation or movement of such cell population.

- 5 Preferably, the protein or peptide has the ability to directly stimulate directed movement of cells. Whether a particular protein has chemotactic activity for a population of cells can be readily determined by employing such protein or peptide in any known assay for cell chemotaxis.

The activity of a protein of the invention may, among other means, be measured by following methods:

- 10 Assays for chemotactic activity (which will identify proteins that induce or prevent chemotaxis) consist of assays that measure the ability of a protein to induce the migration of cells across a membrane as well as the ability of a protein to induce the adhesion of one cell population to another cell population. Suitable assays for movement and adhesion include, without limitation, those described in: CURRENT PROTOCOLS IN IMMUNOLOGY, Coligan *et al.*,
15 eds. (Chapter 6.12, MEASUREMENT OF ALPHA AND BETA CHEMOKINES 6.12.1-6.12.28); Taub *et al. J Clin Invest* 95:1370-1376, 1995; Lind *et al. APMIS* 103:140-146, 1995; Muller *et al., Eur J Immunol* 25: 1744-1748; Gruber *et al. J Immunol* 152:5860-5867, 1994; Johnston *et al., J Immunol* 153: 1762-1768, 1994.

Hemostatic and Thrombolytic Activity

- 20 A protein of the invention may also exhibit hemostatic or thrombolytic activity. As a result, such a protein is expected to be useful in treatment of various coagulation disorders (including hereditary disorders, such as hemophilias) or to enhance coagulation and other hemostatic events in treating wounds resulting from trauma, surgery or other causes. A protein of the invention may also be useful for dissolving or inhibiting formation of thromboses and for
25 treatment and prevention of conditions resulting therefrom (such as, for example, infarction of cardiac and central nervous system vessels (*e.g.*, stroke).

The activity of a protein of the invention may, among other means, be measured by the following methods:

- 30 Assay for hemostatic and thrombolytic activity include, without limitation, those described in: Linet *et al., J. Clin. Pharmacol.* 26:131-140, 1986; Burdick *et al., Thrombosis Res.*

45:413-419, 1987; Humphrey *et al.*, *Fibrinolysis* 5:71-79 (1991); Schaub, *Prostaglandins* 35:467-474, 1988.

Receptor/Ligand Activity

A protein of the present invention may also demonstrate activity as receptors, receptor
5 ligands or inhibitors or agonists of receptor/ligand interactions. Examples of such receptors and
ligands include, without limitation, cytokine receptors and their ligands, receptor kinases and
their ligands, receptor phosphatases and their ligands, receptors involved in cell—cell
interactions and their ligands (including without limitation, cellular adhesion molecules (such as
selectins, integrins and their ligands) and receptor/ligand pairs involved in antigen presentation,
10 antigen recognition and development of cellular and humoral immune responses). Receptors and
ligands are also useful for screening of potential peptide or small molecule inhibitors of the
relevant receptor/ligand interaction. A protein of the present invention (including, without
limitation, fragments of receptors and ligands) may themselves be useful as inhibitors of
receptor/ligand interactions.

15 The activity of a protein of the invention may, among other means, be measured by the
following methods:

Suitable assays for receptor-ligand activity include without limitation those described in:
CURRENT PROTOCOLS IN IMMUNOLOGY, Ed by Coligan, *et al.*, Greene Publishing Associates and
Wiley-Interscience (Chapter 7.28, Measurement of Cellular Adhesion under static conditions
20 7.28.1-7.28.22), Takai *et al.*, *Proc Natl Acad Sci USA* 84:6864-6868, 1987; Bierer *et al.*, *J. Exp.*
Med. 168:1145-1156, 1988; Rosenstein *et al.*, *J. Exp. Med.* 169:149-160 1989; Stoltenborg *et al.*,
J Immunol Methods 175:59-68, 1994; Stitt *et al.*, *Cell* 80:661-670, 1995.

Anti-Inflammatory Activity

Proteins of the present invention may also exhibit anti-inflammatory activity. The
25 anti-inflammatory activity may be achieved by providing a stimulus to cells involved in the
inflammatory response, by inhibiting or promoting cell—cell interactions (such as, for example,
cell adhesion), by inhibiting or promoting chemotaxis of cells involved in the inflammatory
process, inhibiting or promoting cell extravasation, or by stimulating or suppressing production
of other factors which more directly inhibit or promote an inflammatory response. Proteins
30 exhibiting such activities can be used to treat inflammatory conditions including chronic or acute

conditions), including without limitation inflammation associated with infection (such as septic shock, sepsis or systemic inflammatory response syndrome (SIRS)), ischemia-reperfusion injury, endotoxin lethality, arthritis, complement-mediated hyperacute rejection, nephritis, cytokine or chemokine-induced lung injury, inflammatory bowel disease, Crohn's disease or resulting from
5 over production of cytokines such as TNF or IL-1. Proteins of the invention may also be useful to treat anaphylaxis and hypersensitivity to an antigenic substance or material.

Tumor Inhibition Activity

In addition to the activities described above for immunological treatment or prevention of tumors, a protein of the invention may exhibit other anti-tumor activities. A protein may inhibit
10 tumor growth directly or indirectly (such as, for example, via ADCC). A protein may exhibit its tumor inhibitory activity by acting on tumor tissue or tumor precursor tissue, by inhibiting formation of tissues necessary to support tumor growth (such as, for example, by inhibiting angiogenesis), by causing production of other factors, agents or cell types which inhibit tumor growth, or by suppressing, eliminating or inhibiting factors, agents or cell types which promote
15 tumor growth.

Other Activities

A protein of the invention may also exhibit one or more of the following additional activities or effects: inhibiting the growth, infection or function of, or killing, infectious agents, including, without limitation, bacteria, viruses, fungi and other parasites; effecting (suppressing
20 or enhancing) bodily characteristics, including, without limitation, height, weight, hair color, eye color, skin, fat to lean ratio or other tissue pigmentation, or organ or body part size or shape (such as, for example, breast augmentation or diminution, change in bone form or shape); effecting biorhythms or circadian cycles or rhythms; effecting the fertility of male or female subjects; effecting the metabolism, catabolism, anabolism, processing, utilization, storage or
25 elimination of dietary fat, lipid, protein, carbohydrate, vitamins, minerals, cofactors or other nutritional factors or component(s); effecting behavioral characteristics, including, without limitation, appetite, libido, stress, cognition (including cognitive disorders), depression (including depressive disorders) and violent behaviors; providing analgesic effects or other pain reducing effects; promoting differentiation and growth of embryonic stem cells in lineages other
30 than hematopoietic lineages; hormonal or endocrine activity; in the case of enzymes, correcting

deficiencies of the enzyme and treating deficiency-related diseases; treatment of hyperproliferative disorders (such as, for example, psoriasis); immunoglobulin-like activity (such as, for example, the ability to bind antigens or complement); and the ability to act as an antigen in a vaccine composition to raise an immune response against such protein or another material or
5 entity which is cross-reactive with such protein.

Neural disorders in general include Parkinson's disease, Alzheimer's disease, Huntington's disease, multiple sclerosis, amyotrophic lateral sclerosis (ALS), peripheral neuropathy, tumors of the nervous system, exposure to neurotoxins, acute brain injury, peripheral nerve trauma or injury, and other neuropathies, epilepsy, and/or tremors.

10 EQUIVALENTS

From the foregoing detailed description of the specific embodiments of the invention, it should be apparent that particular novel compositions and methods involving nucleic acids, polypeptides, antibodies, detection and treatment have been described. Although these particular
15 embodiments have been disclosed herein in detail, this has been done by way of example for purposes of illustration only, and is not intended to be limiting with respect to the scope of the appended claims that follow. In particular, it is contemplated by the inventors that various substitutions, alterations, and modifications may be made as a matter of routine for a person of ordinary skill in the art to the invention without departing from the spirit and scope of the invention as defined by the claims. Indeed, various modifications of the invention in addition to
20 those described herein will become apparent to those skilled in the art from the foregoing description and accompanying figures. Such modifications are intended to fall within the scope of the appended claims.

THIS PAGE INTENTIONALLY LEFT BLANK

THIS PAGE INTENTIONALLY LEFT BLANK

THIS PAGE INTENTIONALLY LEFT BLANK

THIS PAGE INTENTIONALLY LEFT BLANK

THIS PAGE INTENTIONALLY LEFT BLANK

THIS PAGE INTENTIONALLY LEFT BLANK

WO 00/58473

PCT/US00/08621

THIS PAGE INTENTIONALLY LEFT BLANK

THIS PAGE INTENTIONALLY LEFT BLANK

Table 1

ORF#	Internal Identification Number	Protein similarity	Protein domain	Protein Classification	Cells or Tissues in which Gene is Expressed
1	13076366 (1, 2)	Novel Protein sim. GBank gj4691395[emb]CAB41562.1] - (AL049727) putative large secreted protein [Streptomyces coelicolor]		UNCLASSIFIED	264636
2	80248091 (3, 4)	Novel Protein sim. GBank gj2829508[sp]P71559[SUCC_MYCTU - SUCCINYL-COA SYNTHETASE BETA CHAIN (SCS-BETA)]	Contains protein domain (PF00549) - CoA-ligases	UNCLASSIFIED	264807, 264600, 264602, 264762, 264769, 264689, 264638, 264567
3	80415924 (5, 6)			UNCLASSIFIED	264810, 264604, 264634, 264905, 264636, 264691, 264907, 264692, 264829
4	62018837 (7, 8)			UNCLASSIFIED	264908, 264909, 264760, 264628, 264635
5	79870035 (9, 10)			UNCLASSIFIED	22278002, 264563
6	79842462 (11, 12)		Contains protein domain (PF00127) - Copper binding proteins, plastocyanin/azurin family	UNCLASSIFIED	264908
7	85515576 (13, 14)	Novel Protein sim. GBank gj4415926[gb]AAD20157] - (AC006282) unknown protein [Arabidopsis thaliana]		UNCLASSIFIED	20281099, 35696052, 264508, 264509, 264905, 264906, 264907, 264908, 264909, 264511, 265006, 264512, 265009, 264910, 264595, 264596, 264758, 264603, 264804, 264760, 264762, 264683, 264766, 264787, 264689, 35695917, 264890, 264892, 264693, 33857109, 264828, 264629, 35696423, 55811576, 35695855, 264630, 264631, 264632, 264834, 264638, 264637, 264638, 264839, 18108385, 264563, 264564, 264566, 264486
8	56924278 (15, 16)	Novel Protein sim. GBank gj585562[sp]D06458[NIRB_KLEPN - NITRITE REDUCTASE (NAD(P)H) LARGE SUBUNIT]		reductase	264907
9	79394457 (17, 18)			UNCLASSIFIED	265007, 265019, 263972
10	79556459 (19, 20)			UNCLASSIFIED	264908
11	20414027 (21, 22)			UNCLASSIFIED	264605
12	94141210 (23, 24)	Novel Protein sim. GBank gj3878145[emb]CAA99871] - (Z75543) similar to potassium channel protein [Caenorhabditis elegans]		misc_channel	264259, 265007, 83373044
13	20750551 (25, 26)				
14	95105114 (27, 28)	Novel Protein sim. GBank gj2832781[emb]CAA12645] - (A225805) inward potassium channel alpha subunit [Egeria densa]	Contains protein domain (PF00023) - Ank repeat	UNCLASSIFIED	264556, 264557, 264564
15	20458307 (29, 30)	Novel Protein sim. GBank gj1710791[sp]Q10234[FTOS_SCHPO - PROBABLE MITOCHONDRIAL 40S RIBOSOMAL PROTEIN S5]	Contains protein domain (PF00333) - Ribosomal protein S5	potassium_channel	35696288, 35696052, 264510, 35695917, 264691, 264628, 35696423, 264555, 264558, 264559, 83373044
16	20760358 (31, 32)			ribosomalprot	264604
					264555

17	20292744 (33, 34)	Novel Protein sim. GBank gll1174894sp P44391 URE1_HAEN - UREASE ALPHA SUBUNIT (UREA AMIDOHYDROLASE)	Contains protein domain (PF00449) - Urease		264600
18	60246804 (35, 36)	Novel Protein sim. GBank gll2281102 (AC002333) - SF16 Isolog [Arabidopsis thaliana]			29331827, 284555, 284557, 264638, 284558
19	80078624 (37, 38)				
20	20724556 (39, 40)	Novel Protein sim. GBank gll2508112sp P43872 UUP_ECOLI - ABC TRANSPORTER ATP-BINDING PROTEIN UUP	UNCLASSIFIED transport		22278998, 284907, 284910, 284600, 284893 284602
21	80417554 (41, 42)	Novel Protein sim. GBank gll1730203sp P50442 GATM_RAT - GLYCINE AMIDINOTRANSFERASE PRECURSOR (L- ARGININE:GLYCINE AMIDINOTRANSFERASE) (TRANSAMIDINASE) (AT)	UNCLASSIFIED		22278995, 284906, 265006, 265010, 265011, 284602, 284605, 284788, 284688, 21908784, 284891, 18108376, 284638, 18108387, 264486
22	11705858 (43, 44)				
23	80419176 (45, 46)	Novel Protein sim. GBank gll1877329emb CAB07077 - (Z92771) lade25 [Mycobacterium tuberculosis]	Contains protein domain (PF00441) - Acyl-CoA dehydrogenase		264685 264488, 264907, 284909, 284600, 284602, 284803, 284805, 284682, 284788, 32833986, 264636, 284488
24	20291697 (47, 48)				
25	80253774 (49, 50)				
26	60255384 (51, 52)				
27	60235795 (53, 54)	Novel Protein sim. GBank gll4808369emb CAB42783.1 - (AL049841) putative 30S ribosomal protein S14 [Streptomyces coelicolor]	UNCLASSIFIED ribosomal prot		22278996, 56182435, 265018, 284558 18108370, 35896423, 284835, 284555
28	79483581 (55, 56)				
29	62448765 (57, 58)	Novel Protein sim. GBank gll3122290sp Q06333 K6PF_STRCO - 8- PHOSPHOFRUCTOKINASE (PHOSPHOFRUCTOKINASE) (PHOSPHOHEXOKINASE) (ATP-PFK)	UNCLASSIFIED kinase		264638 284601, 284782, 284766, 284769, 264638
30	79199333 (59, 60)				
31	19848158 (61, 62)				
32	82448485 (63, 64)	Novel Protein sim. GBank gll3560504 (AF027770) - unknown [Mycobacterium smegmatis]	UNCLASSIFIED		264808, 265019, 264887, 21908764, 21906768 284634
33	79582628 (65, 66)	Novel Protein sim. GBank gll2129003 p gll364507 - hypothetical protein MJ1885 - Methanococcus jannaschii	UNCLASSIFIED		284805, 284605, 284762, 284788, 284687, 264689
34	87467667 (67, 68)				
35	95005170 (69, 70)	Novel Protein sim. GBank gll5420387emb CAB46679.1 - (A1243459) proteophosphoglycan [Leishmania major]	UNCLASSIFIED		60432289, 284600, 284602, 284760, 18108357, 284769, 285020, 284691 284600, 284887, 284556, 284639
36	19642042 (71, 72)	Novel Protein sim. GBank gll3287739sp P73538 BIOB_SYNY3 - BIOTIN SYNTHASE (BIOTIN SYNTHETASE)	synthase		264566
37	20389215 (73, 74)	Novel Protein sim. GBank gll231314 gbl AAC07126.1 - (AE000527) delta-1-pyrroline-5-carboxylate dehydrogenase [Helicobacter pylori 26895]	dehydrogenase		284603

38	20466334 (75, 76)	Novel Protein sim. GBank gl 3805970 emb CAA06231 - (A004933) periplasmic nitrate reductase, large subunit [Rhodospseudomonas sp.]		reductase	264605
39	94300715 (77, 78)	Novel Protein sim. GBank gl 1929449 (L63543) - endodermis [Xenopus laevis]	Contains protein domain (PF00207) - Alpha-2-macroglobulin family	complement	264905, 264906, 264907, 66712502, 264908, 264909, 264511, 265008, 264910, 55812038, 264758, 265011, 264762, 264692, 264763, 264764, 264786, 265022, 264893, 264628, 264631, 264634, 264635, 264555, 264638, 16106361, 264558, 16106385, 264462
40	20535625 (78, 80)	Novel Protein sim. GBank gl 854065 emb CAA58337 - (X83413) U68 [Human herpesvirus 6]		UNCLASSIFIED	264592
41	80023287 (81, 82)	Novel Protein sim. GBank gl 854065 emb CAA58337 - (X83413) U68 [Human herpesvirus 6]			264591, 35695917
42	20724566 (83, 84)	Novel Protein sim. GBank gl 3820584 (AF086791) - carbamoylphosphate synthetase large subunit [Zymomonas mobilis]		UNCLASSIFIED	264602
43	20467068 (85, 86)	Novel Protein sim. GBank gl 3820584 (AF086791) - carbamoylphosphate synthetase large subunit [Zymomonas mobilis]		synthase	264605
44	13085297 (87, 88)	Novel Protein sim. GBank gl 2494784 sp Q50729 GUAA_MYCTU - GMP SYNTHASE (GLUTAMINE-HYDROLYZING) (GLUTAMINE AMIDOTRANSFERASE) (GMP SYNTHETASE)	Contains protein domain (PF00958) - GMP synthase C terminal domain	synthase	264769, 264636
45	39384711 (89, 90)	Novel Protein sim. GBank gl 1881738 (U89688) - myosin-I binding protein Acan125 [Acanthamoeba castellanii]		UNCLASSIFIED	264769, 264510, 264508
46	95003398 (91, 92)			ngl	264566
47	11696624 (93, 94)			UNCLASSIFIED	264689
48	78407218 (95, 96)				18108385, 264635, 264828
49	21659644 (97, 98)			UNCLASSIFIED	264603
50	80503986 (99, 100)				264508, 264603, 264769, 264689, 264636, 264558, 264486
51	80255569 (101, 102)	Novel Protein sim. GBank gl 3411177 (AF076240) - MocC [Rhizobium leguminosarum bv. viciae]		UNCLASSIFIED	264593, 16108387
52	78206328 (103, 104)	Novel Protein sim. GBank gl 3914892 sp Q26284 SM41_HEMPU - 41 KD SPICULE MATRIX PROTEIN PRECURSOR (HSM41) (HPSMC)		strud	264634
53	38996970 (105, 106)	Novel Protein sim. GBank gl 3960411 (AC004561) - putative proline-rich protein [Arabidopsis thaliana]		UNCLASSIFIED	264782
54	79570897 (107, 108)	Novel Protein sim. GBank gl 1633572 (U52064) - Herpesvirus salmudi ORF73 homolog [Kaposi's sarcoma-associated herpes-like virus]		UNCLASSIFIED	264630, 264909, 264766
55	80202703 (109, 110)	Novel Protein sim. GBank gl 4321580 sp AAD15785 - (AF050114) alginate lyase [Pseudomonas sp. W7]			29331824, 264102, 265016, 16108376
56	6758408 (111, 112)				264604
57	11223386 (113, 114)		Contains protein domain (PF00076) - RNA recognition motif (a.k.a. FRM, RBD, or RNP domain)	UNCLASSIFIED	264557

58	91227508 (115, 118)	Novel Protein sim. GBank gij5616074 gb AAD45618.1 AF06194. (AF061943) protease- derived STE20-like kinase PSK [Homo sapiens]	Contains protein domain (PF00069) - Eukaryotic protein kinase domain	kinase	56192575, 264259, 80432048, 35688052, 66712502, 264909, 265008, 285010, 285011, 284681, 29148784, 35689917, 60170615, 264891, 264892, 264893, 18108374, 35686423, 58182323, 80432113 264600, 264889, 264638
59	80077371 (117, 118)	Novel Protein sim. GBank gij1172920 sp P45830 REF_MYCLE - PUTATIVE UNDECAPRENYL-PHOSPHATE ALPHA-N- ACETYLGLUCOSAMINYLTRANSFERASE	Contains protein domain (PF00953) - Glycosyl transferase	transferase	
80	12958341 (119, 120)	Novel Protein sim. GBank gij1710216 (U78260) - unknown			
81	80428808 (121, 122)	[Homo sapiens]		glycoprotein	264689 264768
82	13504868 (123, 124)				264630
83	16474553 (125, 126)			UNCLASSIFIED	265019
84	20724578 (127, 128)	Novel Protein sim. GBank gij420845 p jA47041 - transposase homolog (insertion element ISAE1) - Alcaligenes eutrophus		UNCLASSIFIED	264802
85	78328308 (129, 130)	Novel Protein sim. GBank gij3122312 sp O06134 KPYK_MYCTU - PYRUVATE KINASE (PK)	Contains protein domain (PF00224) - Pyruvate kinase	kinase	264563
86	48854384 (131, 132)	Novel Protein sim. GBank gij3928723 emb CAA22218 - (AL034355) putative ABC transporter [Streptomyces coelicolor]		transport	22278896, 264558
87	78952543 (133, 134)	Novel Protein sim. GBank gij231985 sp P30234 DHA_MYCTU - ALANINE DEHYDROGENASE (40 KD ANTIGEN)		dehydrogenase	285021
88	79817382 (135, 136)				264909
89	79841764 (137, 138)			UNCLASSIFIED	264908
90	79871329 (139, 140)				264908, 284908
91	85897458 (141, 142)			UNCLASSIFIED	264602, 265021
92	87734977 (143, 144)	Novel Protein sim. GBank gij4415826 gb AAD20157 - (AC006282) unknown protein [Arabidopsis thaliana] pombg]		UNCLASSIFIED	264488, 264905, 264906, 264807, 264908, 264511, 265008, 264910, 264758, 87168474, 284882, 264766, 264888, 264689, 35695917, 285021, 80170815, 264891, 33637023, 264692, 284893, 264829, 264631, 264639, 22279000
93	80025241 (145, 146)			UNCLASSIFIED	60424178, 264508, 264908, 265007, 264603, 264687, 284889, 264892, 18108387
94	20377410 (147, 148)			UNCLASSIFIED	264605
95	11818032 (149, 150)	Novel Protein sim. GBank gij285309 emb CAA16914 - (AL021787) vacuolar protein sorting [Schizosaccharomyces pombg]		UNCLASSIFIED	284889
96	95105303 (151, 152)	Novel Protein sim. GBank gij4468811 emb CAB38212 - (AL035601) putative protein [Arabidopsis thaliana]		UNCLASSIFIED	83373044, 264906, 264557
97	10144718 (153, 154)	Novel Protein sim. GBank gij854085 emb CAA58337 - (X83413) U88 [Human herpesvirus 8]		UNCLASSIFIED	284583
98	87582566 (155, 156)			UNCLASSIFIED	264804

79	94140190 (157, 158)	Novel Protein sim. GBank gi 5689453 dbj BAA83010.1 - (AB028981) KIAA1056 protein [Homo sapiens]	Contains protein domain (PF00189) - PH domain		35698288, 22278998, 29331822, 29331824, 29331825, 29331827, 284908, 284909, 285008, 284907, 86712502, 284908, 284909, 285008, 285009, 284910, 60170831, 55812038, 33109954, 285017, 285018, 284288, 284768, 58161582, 21906765, 21906769, 29148784, 285020, 284890, 284891, 284892, 284893, 60431528, 35698423, 284831, 284832, 284834, 284836, 284839, 83373044, 284584, 284586, 284587
80	82314840 (159, 160)		UNCLASSIFIED		284769, 284601, 285006, 284910, 284604, 284605, 284634, 284635, 284905, 284782, 284637, 284592, 284626, 284907, 284691, 284908, 284587, 284909, 284768
81	20467247 (161, 162)	Novel Protein sim. GBank gi 1723442 sp Q10258 YD2A_SCHPO - HYPOTHETICAL 89.0 KD PROTEIN C56F8.10 IN CHROMOSOME 1	reductase		284605
82	16331388 (163, 164)	Novel Protein sim. GBank gi 2895866 (AF045770) - methylmalonate semi-aldehyde dehydrogenase [Oryza sativa]	dehydrogenase		284587
83	94741160 (165, 166)	Novel Protein sim. GBank gi 3402673 (AC004697) - unknown protein [Arabidopsis thaliana]	UNCLASSIFIED		284488, 284508, 284509, 284805, 284808, 284909, 284511, 284591, 284593, 284594, 284595, 284596, 284758, 284603, 284760, 284681, 18108351, 284762, 284682, 284764, 284684, 284766, 284686, 284632, 284637, 284557, 284638, 284639, 18108385, 284566
84	80355375 (167, 168)	Novel Protein sim. GBank gi 117336 sp P45380 SAT1_RAT - SULFATE ANION TRANSPORTER 1 (CANALICULAR SULFATE TRANSPORTER) (SULFATE/CARBONATE ANTI-PORTER)	transport		284506, 284906, 284907, 284908, 284909, 284910, 284780, 284763, 284784, 284768, 284768, 284769, 35895855, 284638, 284637
85	60499600 (169, 170)	Novel Protein sim. GBank gi 2120998 pf J570682 - glycosyltransferase homolog - Bordetella pertussis	transferase		284605, 284762, 284687, 284769, 18108374, 284638, 284488
86	39559043 (171, 172)	Novel Protein sim. GBank gi 3256023 emb CAA17228.1 - (AL021897) hypothetical protein Rv1112 [Mycobacterium tuberculosis]			284910
87	13856808 (173, 174)		UNCLASSIFIED		264093

88	95344718 (175, 176)	Novel Protein sim. GBank gl 559703 db BAA07552 - (C38548) ha1025 is new [Homo sapiens]		52844507, 52846365, 18108398, 65274572, 56182575, 56994075, 35696286, 22278997, 22278998, 22278999, 264092, 264093, 284094, 264085, 264259, 29331822, 29331824, 56182181, 88714117, 29331825, 29331826, 60432289, 29331827, 29331828, 35696052, 33656970, 284105, 264508, 284905, 284906, 264907, 264908, 29331830, 68712502, 52844045, 58182435, 265007, 265009, 264910, 60170831, 264592, 60433438, 55812038, 264758, 21906754, 52846317, 33108954, 52844298, 87189474, 265011, 87168559, 264601, 265017, 265018, 264604, 285019, 284448, 264369, 264288, 264768, 52844229, 21906786, 21906787, 21906768, 21906769, 55811957, 35695917, 265020, 285021, 265022, 60170815, 52844150, 33657023, 65274820, 33657109, 27488281, 27486264, 33857349, 35695763, 264628, 263972, 18108374, 55810764, 35696423, 55811578, 65274791, 35895855, 60431850, 264638, 52844332, 56182323, 60170394, 83373044, 18108385, 18108387, 18108388, 56526488, 87168518, 60432113, 22279000, 22278002, 264482, 264584, 264486, 264600
89	80077388 (177, 178)	Novel Protein sim. GBank gl 1710383 sp P46352 RIPX_BACSU - PROBABLE INTEGRASE/RECOMBINASE RIPX		
90	82115988 (179, 180)	Novel Protein sim. GBank gl 249889 sp P78403 VEGQ_ECOLI - PUTATIVE PROTEASE IN BAER-OSRK INTERGENIC REGION	UNCLASSIFIED	264780
91	76906950 (181, 182)	Novel Protein sim. GBank gl 3367754 emb CAA20078 - (AL031155) hypothetical protein SC3A7.16c [Sireplomyces coelicolor]	protease	265008
92	79554871 (183, 184)	Novel Protein sim. GBank gl 3367754 emb CAA20078 - (AL031155) hypothetical protein SC3A7.16c [Sireplomyces coelicolor]	UNCLASSIFIED	264681
93	80496778 (185, 186)	Novel Protein sim. GBank gl 285085 (AF011337) - putative E1-E2 ATPase [Mus musculus]		
94	79646849 (187, 188)	Novel Protein sim. GBank gl 171919 sp P46920 OPUA_BACSU - GLYCINE BETAINES TRANSPORT ATP-BINDING PROTEIN OPUBA	ATPase associated	264907, 264908, 264910, 265009, 264605, 264789
95	11080238 (189, 190)		Contains protein domain (PF00571) - CBS domain	264908
				264594

98	94322125 (101, 192)	Novel Protein sim. GBank gij4589580dbj BAA76802.1 - (AB023175) KIAA0956 protein [Homo sapiens]	UNCLASSIFIED	22276995, 22276999, 264259, 29331822, 29331828, 35696052, 29146498, 264509, 264906, 264907, 264908, 264909, 265007, 265008, 264910, 265009, 264593, 265010, 265017, 264604, 265018, 16108351, 264268, 264768, 264768, 264769, 21906765, 21906767, 21906768, 265020, 264692, 33657182, 35695783, 264828, 264629, 16108378, 264631, 264636, 18106381, 264559, 16108382, 83373044, 22278002, 264508
97	79605200 (193, 194)	Novel Protein sim. GBank gij4583559jemb CAB40388.1 - (AJ005255) OxyR [Erwinia chrysanthemi]	UNCLASSIFIED	
98	79427000 (195, 196)	Novel Protein sim. GBank gij1001693dbj BAA10430 - (DS4002) hypothetical protein [Synecocystis sp.]	UNCLASSIFIED	264909
99	20466524 (197, 198)	Novel Protein sim. GBank gij1169478jsep P43925 EFG_HAEIN - ELONGATION FACTOR G (EF-G)	UNCLASSIFIED	264605
100	79640113 (199, 200)	Novel Protein sim. GBank gij480897jpr J537485 - gene msg1 protein - mouse	UNCLASSIFIED	264693
101	80203298 (201, 202)	Novel Protein sim. GBank gij2894166jemb CAA11773.1 - (AJ223998) PCZA381.18 [Amycolatopsis orientalis]	UNCLASSIFIED	265020, 264102, 263972
102	20467259 (203, 204)	Novel Protein sim. GBank gij1731040jsep P54509YQH_H_BACSU - HYPOTHETICAL HELICASE IN SINI-GCVT INTERGENIC REGION	synthase	264605
103	20466388 (205, 206)	Novel Protein sim. GBank gij654065jemb CAA58337 - (X63413) U88 [Human herpesvirus 6]	helicase	264605
104	80247572 (207, 208)	Novel Protein sim. GBank gij1685117 (U70770) - furrowed [Drosophila melanogaster]	UNCLASSIFIED	264591, 264595, 264602
105	79605208 (209, 210)	Novel Protein sim. GBank gij1705505jsep P54729 BS4_MOUSE - BS4 PROTEIN	complement	264508
106	26362058 (211, 212)	Novel Protein sim. GBank gij4687229jeb AAD32244.1 AF150755 - (AF150755) microtubule-actin crosslinking factor [Mus musculus]	UNCLASSIFIED	264511, 265009
107	80057791 (213, 214)	Novel Protein sim. GBank gij4687229jeb AAD32244.1 AF150755 - (AF150755) microtubule-actin crosslinking factor [Mus musculus]	ATPase-associated	26331824, 264591, 21906754, 265019
108	80237936 (215, 216)	Novel Protein sim. GBank gij263577jemb CAB15284 - (Z99120) similar to ABC transporter (ATP-binding protein) [Bacillus subtilis]	transport	16108374, 35695917, 22278998, 264113, 264600, 264602, 264603, 265017, 264910, 264906, 264638, 264768
109	95194148 (217, 218)	Novel Protein sim. GBank gij233078jemb CAB11265 - (Z98801) carboxypeptidase s precursor [Schizosaccharomyces pombe]	UNCLASSIFIED	264758, 264603, 264630, 264636, 264637
110	79582823 (219, 220)			264687
111	39565458 (221, 222)			264564
112	79656038 (223, 224)		UNCLASSIFIED	264908
113	17958439 (225, 226)		UNCLASSIFIED	265007
114	80502101 (227, 228)		UNCLASSIFIED	264769

115	80251003 (229, 230)	Novel Protein sim. GBank gij2246532 (U93872) - ORF 73, contains large complex repeat CR 73 [Kaposi's sarcoma-associated herpesvirus]		UNCLASSIFIED	52645158, 52645080, 33855970, 264592, 21906754, 27486284, 18108378, 35898423, 264635, 52844332, 18108382
116	61288688 (231, 232)				264905, 264908, 264907, 264908, 264909, 264910, 264758, 265010, 264783, 264882, 264764, 264768, 264685, 264688, 264768, 264769, 33857023, 264893, 33857109, 264628, 18108374, 264631, 264632, 264634, 264636, 264637, 264638, 264639, 56528486, 264585, 264566
117	78636695 (233, 234)				264639, 264693
118	80222170 (235, 236)		Contains protein domain (PF00170) - bZIP transcription factor	UNCLASSIFIED	263974
119	91013071 (237, 238)	Novel Protein sim. GBank gij732528 (U22327) - alpha2(IV) collagen [Caenorhabditis elegans]		UNCLASSIFIED	22276996, 29331624, 60432288, 265007, 60433436, 264603, 264605, 18108351, 264769, 264689, 265020, 264534, 27486281, 264556, 83373044, 18108385, 264584
120	6756491 (239, 240)	Novel Protein sim. GBank gij2131219 [S50157 - cyclin-dependent kinase chain SRB10 - yeast (Saccharomyces cerevisiae)]		kinase	264603
121	80028153 (241, 242)				264595
122	20457620 (243, 244)	Novel Protein sim. GBank gij2052147 [embj(CAB08137) - (Z94752) ksgA [Mycobacterium tuberculosis]]	Contains protein domain (PF00398) - Ribosomal RNA adenine dimethylases	transferase	264605
123	6756278 (245, 246)				264604
124	79104017 (247, 248)	Novel Protein sim. GBank gij2833385 [Q43134] [UGST_SORBI - GRANULE-BOUND GLYCOGEN (STARCH) SYNTHASE PRECURSOR		synthase	18108394, 18108397, 285008, 265007, 285008, 265010, 265011, 18108355, 18108379, 18108380, 18108384
125	87787986 (249, 250)	Novel Protein sim. GBank gij475542 (U08255) - glutamate receptor delta-1 subunit [Rattus norvegicus]	Contains protein domain (PF00080) - Ligand-gated ion channel	misc_channel	264508, 284908, 265009, 264596, 22276002
126	55701283 (251, 252)	Novel Protein sim. GBank gij510278 [embj(CAB45200.1) - (AL079308) putative transcriptional regulator (Streptomyces coelicolor]			264511
127	20487287 (253, 254)				
128	80248473 (255, 256)	Novel Protein sim. GBank gij130120 [spj23620] [PHOB_PSEAE - PHOSPHATE REGULON TRANSCRIPTIONAL REGULATORY PROTEIN PHOB	Contains protein domain (PF00072) - Response regulator receiver domain	UNCLASSIFIED - phosphatase	264605, 264907, 264909, 264910, 264600, 264601, 264603, 264605, 18108351, 264693, 264557
129	95290543 (257, 258)	Novel Protein sim. GBank gij2506493 [spj38036] [YCSB_ECOLI - HYPOTHETICAL 100.5 KD PROTEIN IN IAP-CYSH INTERGENIC REGION	Contains protein domain (PF00270) - DEAD/DEAH box helicase	UNCLASSIFIED	35898423, 35895855, 264600, 264602, 264603, 264604, 264605, 264506, 264908, 264584, 264828, 264682, 264585, 264683
130	80065563 (259, 260)	Novel Protein sim. GBank gij654085 [embj(CAA58337) - (X63413) U88 [Human herpesvirus 6]			264634
131	84985022 (261, 262)	Novel Protein sim. GBank gij1076038 [prij(S54860) - ABC transporter PatC-2 chain - Mycobacterium tuberculosis]	Contains protein domain (PF00528) - Binding-protein-dependent transport systems inner membrane component	transport	18108376, 264769, 29331828, 264689, 22278986, 265021, 264800, 264511, 264601, 264602, 264605, 264905, 264638

132	10887692 (283, 264)	Novel Protein sim. GBank gll1877340[emb]CAB07068] - (Z52771) eccA3 [Mycobacterium tuberculosis]	Contains protein domain (PF00289) - Carbamoyl-phosphate synthase (CPSase)	carboxylase	264636 264905, 264689
133	94630883 (265, 266)	Novel Protein sim. GBank gll4585638[emb]CAB40932.1] - (AL049630) putative NADH dehydrogenase [Streptomyces coelicolor]		dehydrogenase	264905, 264605, 265021
134	79834860 (267, 268)	Novel Protein sim. GBank gll1460074[emb]CAB01048] - (Z77250) hypothetical protein Rv2568 [Mycobacterium tuberculosis]			264834
135	18885057 (269, 270)	Novel Protein sim. GBank gll2125698[emb]CAA7351] - (Y13070) folypolyglutamate synthase [Streptomyces coelicolor]		synthase	264508
136	79846083 (271, 272)	Novel Protein sim. GBank gll5689912[emb]CAB52075.1] - (AL109732) putative mutase [Streptomyces coelicolor A3(2)]			264693, 264685, 264688, 264691, 264692, 264893
137	79818770 (273, 274)	Novel Protein sim. GBank gll5420387[emb]CAB46678.1] - (AJ243459) proteophosphoglycan [Leishmania major]		UNCLASSIFIED	18108374, 18108385, 33657109, 33657182, 265010, 22278998, 285006, 265007, 265008, 265009, 264693
138	78635871 (275, 276)	Novel Protein sim. GBank gll5689912[emb]CAB52075.1] - (AL109732) putative mutase [Streptomyces coelicolor A3(2)]	Contains protein domain (PF01817) - Chorismate mutase	dehydrogenase	22278998, 285007, 284910, 80433356, 265010, 284802, 284805, 284768, 264688, 264769, 264893, 32833988, 18108374, 18108387
139	66688076 (277, 278)	Novel Protein sim. GBank gll3581916[emb]CAA20855] - (AL031545) mutS family DNA mismatch repair protein [Schizosaccharomyces pombe]		UNCLASSIFIED	264906
140	79825758 (279, 280)	Novel Protein sim. GBank gll3581916[emb]CAA20855] - (AL031545) mutS family DNA mismatch repair protein [Schizosaccharomyces pombe]		nuclease	264600
141	20700084 (281, 282)	Novel Protein sim. GBank gll3581916[emb]CAA20855] - (AL031545) mutS family DNA mismatch repair protein [Schizosaccharomyces pombe]			264602, 265017
142	60028104 (283, 284)	Novel Protein sim. GBank gll3581916[emb]CAA20855] - (AL031545) mutS family DNA mismatch repair protein [Schizosaccharomyces pombe]		UNCLASSIFIED	264600
143	11072274 (285, 286)	Novel Protein sim. GBank gll3581916[emb]CAA20855] - (AL031545) mutS family DNA mismatch repair protein [Schizosaccharomyces pombe]			263976, 264600, 264910, 264632, 264508, 264563, 284564, 284591, 284558, 284908, 284829, 284839
144	95008102 (287, 288)	Novel Protein sim. GBank gll3581916[emb]CAA20855] - (AL031545) mutS family DNA mismatch repair protein [Schizosaccharomyces pombe]		UNCLASSIFIED	22278998, 264602
145	60027058 (289, 290)	Novel Protein sim. GBank gll3581916[emb]CAA20855] - (AL031545) mutS family DNA mismatch repair protein [Schizosaccharomyces pombe]			
146	13065662 (291, 292)	Novel Protein sim. GBank gll140807[sp]P24536[YI21_BURCE - INSERTION ELEMENT IS402 HYPOTHETICAL 24 KD PROTEIN]	Contains protein domain (PF01875) - Transposase		264687
147	94320368 (293, 294)	Novel Protein sim. GBank gll2827608[emb]CAA16663] - (AL021646) uvrD2 [Mycobacterium tuberculosis]		helicase	264905, 284906, 264909, 284510, 265008, 60433358, 264600, 264601, 264604, 264605, 264687, 264769, 18108385, 85274791, 18108387
148	80248804 (295, 296)	Novel Protein sim. GBank gll2816947[emb]CAA17565] - (AL021999) hypothetical protein Rv0988 [Mycobacterium tuberculosis]		transport	265009, 265010, 264600, 264602, 264803, 264804, 264605, 264693, 33657109, 264636

149	80249373 (297, 298)	Novel Protein sim. GBank gll1723073sp Q11040 Y081_MYCTU - HYPOTHETICAL ABC TRANSPORTER ATP-BINDING PROTEIN CY50.01	Contains protein domain ABC transporter	transport	265010, 264800, 264801, 264603, 264604, 27488285, 264836
150	20284748 (289, 300)	Novel Protein sim. GBank gll3724125 emb CAA11805 - (A)224340) maltose phosphorylase [<i>Lactobacillus</i> sanfrancisco]			264800
151	20726398 (301, 302)	Novel Protein sim. GBank gll729312sp P07851 DEOB_ECOLI - PHOSPHOPENTOMUTASE (PHOSPHODEOXYRIBOMUTASE)	Contains protein domain (PF01676) - Metalloenzyme superfamily	UNCLASSIFIED	264602
152	95002877 (303, 304)	Novel Protein sim. GBank gll2497852sp P55667 Y4TM_RHSN - HYPOTHETICAL HYDROLASE/PEPTIDASE Y4TM		peptidase	264602
153	80256665 (305, 306)	Novel Protein sim. GBank gll3123021sp Q80508 VT1_FUNHE - VITELLOGENIN I PRECURSOR (VTG I) (CONTAINS: LIPOVITELLIN 1 (LV1); PHOSVITIN (PV); LIPOVITELLIN 2 (LV2))		UNCLASSIFIED	264593
154	82305966 (307, 308)	Novel Protein sim. GBank gll419697 p JN0443 - transcription initiation factor sigma homolog hrdB - <i>Streptomyces aureofaciens</i>	Contains protein domain (PF00140) - Sigma-70 factor	mapolymerase	264910, 264782, 264891, 264634 264605
155	20429859 (309, 310)	Novel Protein sim. GBank gll628710 p J S41739 - hypothetical protein - <i>Escherichia coli</i>		UNCLASSIFIED	264565
156	39564742 (311, 312)	Novel Protein sim. GBank gll3695013 (AF052586) - CIRA [<i>Pseudomonas aeruginosa</i>]	Contains protein domain (PF00142) - 4Fe-4S iron sulfur cluster binding proteins, NifH/HfxC family	hydrolase	264691
157	10358887 (313, 314)	Novel Protein sim. GBank gll1073072 p J C55543 - cmaU proteoh - <i>Pseudomonas syringae</i> pv. <i>syringae</i>		UNCLASSIFIED	264905
158	78761838 (315, 316)	Novel Protein sim. GBank gll1073072 p J C55543 - cmaU proteoh - <i>Pseudomonas syringae</i> pv. <i>syringae</i>		UNCLASSIFIED	265008 264605
159	78890378 (317, 318)	Novel Protein sim. GBank gll1173023sp P48789 RL30_STRCO - 50S RIBOSOMAL PROTEIN L30	Contains protein domain (PF00400) - WD domain, G-beta repeat		
160	11075119 (319, 320)	Novel Protein sim. GBank gll1173023sp P48789 RL30_STRCO - 50S RIBOSOMAL PROTEIN L30	Contains protein domain (PF00327) - Ribosomal protein L30pL7e	ribosomelprol	22278996, 264600, 264603, 35695917, 32833986, 35696423, 264638
161	80055007 (321, 322)	Novel Protein sim. GBank gll5304889 emb CAB46028.1 - (AL031685) dJ963K23.2 (novel protein) [Homo sapiens]	Contains protein domain (PF00097) - Zinc finger, C3HC4 type (RING finger)	Interleukin	264112, 264532, 22279002
162	80016371 (323, 324)	Novel Protein sim. GBank gll5304889 emb CAB46028.1 - (AL031685) dJ963K23.2 (novel protein) [Homo sapiens]			
163	11892308 (325, 326)	Novel Protein sim. GBank gll2681681 emb CAA15785 - (AL005204) putative protease [<i>Streptomyces coelicolor</i>]		UNCLASSIFIED	264639
164	80077802 (327, 328)	Novel Protein sim. GBank gll4418478 p J AAD20378 - (AF125999) transposase [<i>Mycobacterium avium</i>]		UNCLASSIFIED	264905, 264907, 264600
165	10856067 (329, 330)	Novel Protein sim. GBank gll4418478 p J AAD20378 - (AF125999) transposase [<i>Mycobacterium avium</i>]		UNCLASSIFIED	264891
166	88095003 (331, 332)	Novel Protein sim. GBank gll76177 p J QOECFT - hypothetical 38.8K protein (fla 5' region) - <i>Escherichia coli</i>		UNCLASSIFIED	264605, 264486
167	16395460 (333, 334)	Novel Protein sim. GBank gll76177 p J QOECFT - hypothetical 38.8K protein (fla 5' region) - <i>Escherichia coli</i>		UNCLASSIFIED	265010
168	80079362 (335, 336)	Novel Protein sim. GBank gll76177 p J QOECFT - hypothetical 38.8K protein (fla 5' region) - <i>Escherichia coli</i>			264800
169	80239581 (337, 338)	Novel Protein sim. GBank gll76177 p J QOECFT - hypothetical 38.8K protein (fla 5' region) - <i>Escherichia coli</i>			264556, 264557, 264558, 264559

170	79612384 (338, 340)	Novel Protein sim. GBank	Contains protein domain (PF01810) - LysE type translocator	284988	
171	95283073 (341, 342)	g1140888 sp P27847 YIGK_ECOLI - HYPOTHETICAL 15.4 KD PROTEIN IN RECO-PLDB INTERGENIC REGION (F138)		284595, 284604	
172	37797007 (343, 344)	Novel Protein sim. GBank g11421095 gb AAD12048.1 - (AF045609) AgIG [Sinorhizobium meliloti]	Contains protein domain (PF00528) - Binding-protein-dependent transport systems inner membrane component	284769	
173	57528660 (345, 346)	Novel Protein sim. GBank g1132854 sp P02387 RL2_ECOLI - 50S RIBOSOMAL PROTEIN L2	Contains protein domain (PF00161) - Ribosomal Proteins L2	264769	
174	95293076 (347, 348)	Novel Protein sim. GBank g11861350 db BAA19377 - (AB001488) PROBABLE TRANSPORT PROTEIN, SIMILAR TO ANTIBIOTIC TRANSPORT-ASSOCIATED PROTEIN ACTIII IN STREPTOMYCES COELICOLOR. [Bacillus subtilis]	transport	264510, 264593, 284602, 284803, 284605, 284782, 284893	
175	79758270 (349, 350)	Novel Protein sim. GBank g12072722 amb CAB08328 - (Z95121) manA [Mycobacterium tuberculosis]	isomerase	264565	
176	80066896 (351, 352)	Novel Protein sim. GBank g11055188 (U40187) - similar to PIR:AA11724 chicken LD (limb deformity) gene product and to formin; also P-rich region similar to collagen [Caenorhabditis elegans]	UNCLASSIFIED	284907, 264910, 264661, 264558	
177	86584852 (353, 354)	Novel Protein sim. GBank g12326738 emb CAB10952 - (Z98268) hypothetical protein Rv1695 [Mycobacterium tuberculosis]	Contains protein domain (PF01513) - Domain of unknown function	264766, 60424179, 264687, 264698, 264769, 29331628, 60432289, 18108376, 264869, 18108387, 32833986, 22278996, 265020, 284600, 284601, 284602, 284603, 284604, 264805, 264835, 264762, 264636, 264808, 264584, 264637, 264636, 284488, 60433356, 264766	
178	78559528 (355, 356)	Novel Protein sim. GBank g11905596 (U81768) - kinesin-73 [Drosophila melanogaster]	strut	264853, 33657109, 264635	
179	20283112 (357, 358)	Novel Protein sim. GBank	UNCLASSIFIED	264563	
180	80488958 (359, 360)	g11189387 sp P45256 DNAB_HAEIN - REPLICATIVE DNA HELICASE	helicase	264769	
181	78585369 (361, 362)	Novel Protein sim. GBank g13170615 (AF059485) - DOCA [Mus musculus]	UNCLASSIFIED	21806787, 284835, 264639, 18108384	
182	80577899 (363, 364)				
183	11614017 (365, 366)	Novel Protein sim. GBank g11076627 pir S54172 - inorganic pyrophosphatase (EC 3.6.1.1) - common tobacco	UNCLASSIFIED	284259, 35696052, 56182435, 284511, 285018, 33657109, 264555, 264588, 264890	
184	10174167 (367, 368)	Novel Protein sim. GBank g14371280 gb AAD18136 - (AC008280) hypothetical protein [Arabidopsis thaliana]	UNCLASSIFIED	264510	

185	21660822 (368, 370)	Novel Protein sim. GBank gij3008178[emb]CAA18398.1] - (AL022304) putative rna transport regulator [Schizosaccharomyces pombe]		UNCLASSIFIED	264604
186	80070328 (371, 372)	Novel Protein sim. GBank gij28802[sp]P94408[YCLF_BACSU - HYPOTHETICAL 53.3 KD PROTEIN IN SFP-GERKA INTERGENIC REGION]		transport	264595
187	80186611 (373, 374)	Novel Protein sim. GBank gij3150260[emb]CAA18179] - (AL023834) cyclin [Schizosaccharomyces pombe]		UNCLASSIFIED	264389
188	20464942 (375, 376)	Novel Protein sim. GBank gij2145853[pi]jS72938 - nix protein - Mycobacterium leprae		kinase	264805
189	82338215 (377, 378)	Novel Protein sim. GBank gij1581244[d]BAA19271] - (AB001488) SIMILAR TO PYRUVATE OXIDASE AND ACETOLACTATE SYNTHASE. [Bacillus subtilis]		UNCLASSIFIED	35696052, 264802, 264605, 264782, 264688, 35695917, 18108370, 18108372, 264838, 264585
190	80086821 (378, 380)	Novel Protein sim. GBank gij1581244[d]BAA19271] - (AB001488) SIMILAR TO PYRUVATE OXIDASE AND ACETOLACTATE SYNTHASE. [Bacillus subtilis]	Contains protein domain (PF00205) - Thiamine pyrophosphate enzymes	synthase	264583
191	88095012 (381, 382)	Novel Protein sim. GBank gij120228[sp]P28725[FKBP_STRCH - FK506-BINDING PROTEIN (PEPTIDYL-PROLYL CIS-TRANS ISOMERASE) (PPIASE) (ROTAMASE)]	Contains protein domain (PF00254) - FKBP-type peptidyl-prolyl cis-trans isomerases	isomerase	264508, 264804, 264605, 264769, 264555
192	16333379 (383, 384)				264567
193	78910127 (385, 386)				264908, 264893
194	20464949 (387, 388)				264805
195	13510389 (389, 390)	Novel Protein sim. GBank gij4980892[gb]AAD35474.1[AE00171] - (AE001718) ABC transporter. ATP-binding protein [Thermotoga maritima]		transport	264638
196	95005588 (391, 392)	Novel Protein sim. GBank gij1705461[sp]P53656[BIOA_ERWHE - ADENOSYLMETHIONINE-8-AMINO-7-OXONONANOATE AMINOTRANSFERASE (7,8-DIAMINO-PELAGONIC ACID AMINOTRANSFERASE) (DAPA AMINOTRANSFERASE)]	Contains protein domain (PF00202) - gamma Aminotransferases class-II pyridoxal phosphate	gaba	264600, 264689, 264638
197	60248685 (393, 394)	Novel Protein sim. GBank gij3122305[sp]Q2776[KGPF_SCHMA - 6- PHOSPHOFRUCTOKINASE (PHOSPHOFRUCTOKINASE)]	Contains protein domain (PF00365) - kinase	kinase	264802, 264682, 264692, 18108374
198	78163835 (395, 396)				264638
199	78890715 (397, 398)	Novel Protein sim. GBank gij1781203[emb]CAB08110] - (Z63859) gnd [Mycobacterium tuberculosis]	Contains protein domain (PF00393) - 6-phosphogluconate dehydrogenases		265008
200	78413849 (399, 400)	Novel Protein sim. GBank gij2842222 (AF030885) - telomere-associated recQ-like helicase [Ustilago maydis]		UNCLASSIFIED	264595, 264598
201	86945924 (401, 402)	Novel Protein sim. GBank gij2894378[emb]CAA74911.1] - (Y14573) ring finger protein [Hordeum vulgare]		UNCLASSIFIED	29331828, 265007, 264512, 33857402, 264596, 265017, 18108351, 264682, 264883, 264767, 264829, 55810764, 264634, 264835, 58182323, 60432113, 22279000

202	7858046 (403, 404)	Novel Protein sim. GBank gi 231772 sp P30598 CHS1_USTMA - CHITIN SYNTHASE 1 [CHITIN-UDP ACETYL-GLUCOSAMINYL TRANSFERASE 1]	Contains protein domain (PF01644) - Chitin synthase	284600	
203	79843927 (405, 406)	Novel Protein sim. GBank gi 1504042 gb BAA13220 - (D86984) similar to yeast adenylate cyclase (S56776) [Homo sapiens]		22278995, 28331822, 28331825, 28331827, 264908, 21908754, 284863, 21908768, 21906769, 35896423, 284558	
204	78655168 (407, 408)	Novel Protein sim. GBank gi 263300 emb CAB13310 - (Z99111) similar to hypothetical proteins [Bacillus subtilis]	UNCLASSIFIED	284909	transport
205	10080563 (409, 410)	Novel Protein sim. GBank gi 263300 emb CAB13310 - (Z99111) similar to hypothetical proteins [Bacillus subtilis]	UNCLASSIFIED	284604	
206	6756473 (411, 412)	Novel Protein sim. GBank gi 213438 ipr S60678 - polybromo 1 protein - chicken	UNCLASSIFIED	284558	
207	20754522 (413, 414)	Novel Protein sim. GBank gi 213438 ipr S60678 - polybromo 1 protein - chicken		284605	
208	20289281 (415, 416)	Novel Protein sim. GBank		284605, 284689	
209	80071068 (417, 418)	Novel Protein sim. GBank gi 2501040 sp P005814 SYP_MYCTU - PROLYL-TRNA SYNTHETASE (PROLINE-TRNA LIGASE) (PRORS)		284805, 284807, 284909, 284766, 284687, 284691, 284629, 18106374, 284638	
210	80188600 (419, 420)			263978	
211	80034538 (421, 422)	Novel Protein sim. GBank	UNCLASSIFIED	264508, 264905, 264906, 284907, 284808, 284600, 284782, 284534, 284632, 284634, 284635, 284639, 284488	
212	82442474 (423, 424)	Novel Protein sim. GBank gi 5031609 ref NP_005536.1 p SLR - Immunoglobulin superfamily containing leucine-rich repeat		22278998, 264508, 284600, 284602, 284603, 284605, 33657023, 284565, 284466	
213	80248562 (425, 426)	Novel Protein sim. GBank gi 3123359 sp P033123 LEU2 MYCLE - 3- ISOPROPYLMALATE DEHYDRATASE LARGE SUBUNIT (ISOPROPYLMALATE ISOMERASE) (ALPHA-IPM ISOMERASE) (IPMI)	isomerase		
214	80079381 (427, 428)	Novel Protein sim. GBank gi 116236 sp P19421 CH80_COXBU - 60 KO CHAPERONIN (PROTEIN CPN60) (GROEL PROTEIN) (HEAT SHOCK PROTEIN B)	Contains protein domain (PF00116) - TCP-1/cpn60 chaperonin family	264600, 284693	eph
215	14973283 (429, 430)	Novel Protein sim. GBank gi 3417287 (AC002310) - Unknown gene product [Homo sapiens]	UNCLASSIFIED	284629	
216	80177716 (431, 432)	Novel Protein sim. GBank	Contains protein domain (PF00098) - Zinc finger, C2H2 type	284446	
217	78603634 (433, 434)	Novel Protein sim. GBank gi 2508924 sp P49754 VP41_HUMAN - VACUOLAR ASSEMBLY PROTEIN VPS41 HOMOLOG (S53)		264508	
218	80258475 (435, 436)	Novel Protein sim. GBank gi 1173268 sp P38106 RSEA_ECOLI - SIGMA-E FACTOR NEGATIVE REGULATORY PROTEIN	mapolymerase	284594	
219	20438797 (437, 438)	Novel Protein sim. GBank gi 1173268 sp P38106 RSEA_ECOLI - SIGMA-E FACTOR NEGATIVE REGULATORY PROTEIN		284604	
220	13489572 (439, 440)	Novel Protein sim. GBank gi 2884703 (AF052427) - (Z633664) gi B [Mycobacterium tuberculosis]	synthase	284689	nucleoside
221	11267486 (441, 442)	Novel Protein sim. GBank gi 4587313 gb BAA76709.1 - unknown [Trypanosoma cruzi] (AB025248) alpha-1,2-mannosidase [Bacillus sp. M-90]	UNCLASSIFIED	284555	

222	78662602 (443, 444)	Novel Protein sim. GBank gl 1877268 emb CAB07049 - (Z82770) hypothetical protein Rv0143c [Mycobacterium tuberculosis]		UNCLASSIFIED	264605, 264769, 35696423
223	83053669 (445, 446)			UNCLASSIFIED	264908, 264907, 264603
224	79557920 (447, 448)			UNCLASSIFIED	264664, 264893
225	79559541 (449, 450)	Novel Protein sim. GBank gl 2274851 dbj BA21515 - (D84159) 3-7 gene product [Homo sapiens]		UNCLASSIFIED	264692
226	79172397 (451, 452)	Novel Protein sim. GBank gl 888245 (U29468) - C56C10.7 gene product [Caenorhabditis elegans]		UNCLASSIFIED	22278998, 264112, 33657023, 263981
227	81777186 (453, 454)			UNCLASSIFIED	35695917, 264636, 264907
228	79672285 (455, 456)			UNCLASSIFIED	264768, 264907, 264908, 264692, 264593, 264639
229	79838266 (457, 458)			UNCLASSIFIED	264908, 264910
230	11013208 (459, 460)			UNCLASSIFIED	264631
231	20622207 (461, 462)	Novel Protein sim. GBank gl 1635114 emb CAA71733 - (Y10744) homoserine O-acetyltransferase [Leptospira meyeri]		UNCLASSIFIED	264908, 264600, 264603, 264692
232	80055035 (463, 464)			UNCLASSIFIED	264600, 264603, 264605, 264887, 264769
233	80063054 (465, 466)	Novel Protein sim. GBank gl 2642340 (AF032970) - Indolizone propionate hydrolase [Pseudomonas putida]	Contains protein domain (PF00449) - Urease	UNCLASSIFIED	264604
234	7523996 (467, 468)	Novel Protein sim. GBank gl 3510505 (AF030861) - polypeptide [Fugu rubripes]		UNCLASSIFIED	264369
235	80203671 (469, 470)			UNCLASSIFIED	264106
236	78940001 (471, 472)	Novel Protein sim. GBank gl 2104808 emb CAB08605 - (Z85396) PexA [Mycobacterium leprae]		UNCLASSIFIED	264905
237	11755273 (473, 474)			UNCLASSIFIED	264681
238	78461401 (475, 476)			UNCLASSIFIED	264639
239	82435190 (477, 478)	Novel Protein sim. GBank gl 2495617 sp Q57252 YDU_HAEIN - HYPOTHETICAL PROTEIN H1163	Contains protein domain (PF00037) - 4Fe-4S ferredoxins and related iron-sulfur cluster binding domains.	UNCLASSIFIED	264908, 265010, 264603, 264782, 264682, 264638, 264836, 264486
240	21635575 (479, 480)	Novel Protein sim. GBank gl 3183458 sp P75788 YLA_ECOLI - HYPOTHETICAL ABC TRANSPORTER ATP-BINDING PROTEIN YLA	transport	UNCLASSIFIED	264259, 264789
241	80377307 (481, 482)	Novel Protein sim. GBank gl 3675920 emb CAB04111 - (Z81503) predicted using GeneFINDER, similar to collagen; cDNA EST EMBL:D65450 comes from this gene; cDNA EST EMBL:D68886 comes from this gene [Caenorhabditis elegans]	UNCLASSIFIED	UNCLASSIFIED	264908, 264809, 264764, 264639
242	62146454 (483, 484)			UNCLASSIFIED	264489, 264807, 264908, 264511, 264760, 264764, 264692, 264635, 264637
243	79632207 (485, 486)			UNCLASSIFIED	264908
244	80248662 (487, 488)	Novel Protein sim. GBank gl 2624302 emb CAA15575 - (AL008867) ad [Mycobacterium tuberculosis]	dehydrogenase	UNCLASSIFIED	264600, 264602, 264605, 264769, 264689
245	79863543 (489, 490)	Novel Protein sim. GBank gl 2820625 (AF044499) - vgrE protein [Escherichia coli]		UNCLASSIFIED	264907, 264756
246	79162928 (491, 492)	Novel Protein sim. GBank gl 5420387 emb CAB46679.1 - (AJ243459) proteophosphoglycan [Leishmania major]	Contains protein domain (PF01106) - NFU-like domain		264637, 18108381, 18108387, 264565

247	78673165 (493, 494)	Novel Protein sim. GBank gll1638006[emb]CAB08648] - (Z85982) argB [Mycobacterium tuberculosis]		kinase	284909, 284691, 35696423, 16108387
248	80468863 (495, 498)	Novel Protein sim. GBank gll166574[sp]P42464[ATPB_CORGL - ATP SYNTHASE BETA CHAIN		synthase	35698286, 284907, 284511, 284602, 284768, 284686, 285021, 35695855, 18108385
249	78764845 (487, 488)			UNCLASSIFIED	284907, 284910, 285011, 284762, 284638
250	78619880 (499, 500)				21808788, 284892
251	84359489 (501, 502)			UNCLASSIFIED	52645156, 28331622, 28331824, 52644045, 285016, 21808785, 21808788, 285020, 27486281, 27486285, 35695763, 18108376, 284558, 284559, 284565
252	78737756 (503, 504)	Novel Protein sim. GBank gll3327166[db]BAA31651] - (AB014576) KIAA0676 protein [Homo sapiens]			284685, 284687, 284832
253	20443124 (505, 506)	Novel Protein sim. GBank gll3036880[emb]CAA18513] - (AL022374) putative ATP-dependent DNA helicase [Streptomyces coelicolor]		helicase	284604
254	80027421 (507, 508)	Novel Protein sim. GBank gll3915408[sp]O34981YJMB_BACSU - HYPOTHETICAL SYMPORTER IN COTT-RAPA INTERGENIC REGION		UNCLASSIFIED	284506, 284906, 284602, 284667, 285021, 284488
255	11386315 (509, 510)	Novel Protein sim. GBank gll1685720[db]BAA04134] - (D17312) diarrheal toxin [Bacillus cereus]		UNCLASSIFIED	284593
256	80028158 (511, 512)	Novel Protein sim. GBank gll465767[sp]P34422[YL31_CAEEL - HYPOTHETICAL 88.0 KD PROTEIN F44B9.1 IN CHROMOSOME III	Contains protein domain (PF00328) - Prolyl oligopeptidase family	peptidase	284602, 284692
257	20289282 (513, 514)	Novel Protein sim. GBank gll172039[sp]P42315[SCOA_BACSU - PROBABLE SUCCINYL-COA:3-KETOACID-COENZYME A TRANSFERASE SUBUNIT A (SUCCINYL COA:3-OXOACID COA-TRANSFERASE)(OXCT A)	Contains protein domain (PF01144) - Coenzyme A transferase	transferase	284605
258	20459484 (515, 516)	Novel Protein sim. GBank gll3127836[emb]CAA18902] - (AL023496) hypothetical protein [Streptomyces coelicolor]		UNCLASSIFIED	284604
259	78910152 (517, 518)			collagen	284681, 284688, 284692
260	20379437 (519, 520)			UNCLASSIFIED	284682, 284556
261	20285683 (521, 522)	Novel Protein sim. GBank gll12376[sp]P2422[HUTH_STRGR - HISTIDINE AMMONIA-LYASE (HISTIDASE)	Contains protein domain (PF00221) - Phenylalanine and histidine ammonia lyases	UNCLASSIFIED	284600
262	80189317 (523, 524)			UNCLASSIFIED	285017, 284369
263	88095045 (525, 526)	Novel Protein sim. GBank gll3924708[emb]CAA84646] - (Z35597) Weak similarity with sea squirt nitrogen precursor protein (blast score 71); cDNA EST EMBL:702056 comes from this gene; cDNA EST EMBL:D78135 comes from this gene; cDNA EST EMBL:D73147 comes from this gene; cDNA EST EMB...		UNCLASSIFIED	284486, 284905, 284808, 284907, 284908, 284909, 284512, 284910, 284758, 284596, 284604, 285016, 284605, 284760, 18108351, 284763, 284764, 284288, 284768, 284768, 284789, 284891, 284692, 284683, 284628, 284634, 284635, 284655, 284638, 284639
264	87370628 (527, 528)	Novel Protein sim. GBank gll3043734[db]BAA25531] - (AB011177) KIAA0805 protein [Homo sapiens]	Contains protein domain (PF00047) - Immunoglobulin domain	protease	284259, 284908, 21806754, 285016, 285018, 285020

285	95355048 (528, 530)	Novel Protein sim. GBank gl458962jdbj[BAA76834.1] - (A8023207) KIAA0990 protein [Homo sapiens]		kinase	284488, 35896286, 29331824, 56182181, 35896032, 284508, 284905, 264908, 284907, 66712502, 284808, 284809, 284511, 284512, 284910, 284592, 284595, 284758, 264598, 55811388, 284600, 285017, 284803, 284604, 264805, 264760, 18108351, 284782, 284681, 284784, 284288, 284766, 264788, 264769, 21908765, 21908787, 21908789, 285020, 284891, 33657023, 33857109, 33657182, 284628, 35898423, 35895855, 284830, 284631, 284832, 284634, 284635, 284638, 264555, 284838, 83373044, 58528488, 87188516, 284564, 284586, 284488, 264800
266	79588075 (531, 532)			UNCLASSIFIED	284828
287	11382222 (533, 534)			UNCLASSIFIED	284687, 284768, 284689
288	78909568 (535, 536)			UNCLASSIFIED	284602
289	80025810 (537, 538)			UNCLASSIFIED	284693
270	84381144 (539, 540)	Novel Protein sim. GBank gl4507387[re]NP_003182.1[pTARS - threonyl-tRNA synthetase]		UNCLASSIFIED	284908, 284693
271	79552301 (541, 542)	Novel Protein sim. GBank gl4980738[gb]AAD35331.1[AE001707] glucose-1-phosphate adenylyltransferase [Thermotoga maritima]		synthase	284908
272	8674778 (543, 544)	Novel Protein sim. GBank gl1168224[sp]P44569[SNITD_HAEIN - PROBABLE 5' NUCLEOTIDASE PRECURSOR]		UNCLASSIFIED	284688
273	12840694 (545, 546)	Novel Protein sim. GBank gl1168224[sp]P44569[SNITD_HAEIN - PROBABLE 5' NUCLEOTIDASE PRECURSOR]		UNCLASSIFIED	284564
274	39524248 (547, 548)	Novel Protein sim. GBank gl3253159 [AF005355] - Translation initiation factor eIF2C [Oryzopsis cuniculatus]		UNCLASSIFIED	284807, 284808, 264808, 264766, 284788, 284691, 264632, 284638
275	82787041 (548, 550)	Novel Protein sim. GBank gl134920[sp]P21997[SSGP_VOLCA - SULFATED SURFACE GLYCOPROTEIN 185 (SSG 185)]		UNCLASSIFIED	285008, 80432229
276	86871073 (551, 552)	Novel Protein sim. GBank gl128021[sp]P20964[OBG_BACSU - SPOOB-ASSOCIATED GTP-BINDING PROTEIN]		ribosomal prot	284600, 18108387
277	80078735 (553, 554)	Novel Protein sim. GBank gl79839[pir]S03812 - IYVB protein - Micrococcus luteus		UNCLASSIFIED	284689
278	12866947 (555, 556)	Novel Protein sim. GBank gl79839[pir]S03812 - IYVB protein - Micrococcus luteus		nuclease	284508, 284604, 21908784, 284638, 284557, 284404
279	95282719 (557, 558)	Novel Protein sim. GBank gl313180[sp]PQ16964[YLN2_CAEEL - HYPOTHETICAL 46.2 KD TRP-ASP REPEATS CONTAINING PROTEIN D2013.2 IN CHROMOSOME II]			284259
280	5603617 (559, 560)				18108392, 284634, 284555, 284556, 284557, 284558
281	80249589 (561, 562)				
282	18598682 (563, 564)			UNCLASSIFIED	285019
283	20814211 (565, 566)			UNCLASSIFIED	284555

284	91212160 (567, 568)	Novel Protein sim. GBank gij2428094 (U56632) - acetyl xylan esterase; AxeA (Thermotoga neapolitana)	Contains protein domain (PF00300) - Phosphoglycerate mutase family	UNCLASSIFIED	35696052, 28331628, 264508, 284905, 284600, 284602, 284605, 284662, 284764, 56181562, 21908784, 18108376, 284636, 284559, 18108387
285	6757040 (569, 570)			UNCLASSIFIED	284603
286	80503235 (571, 572)	Novel Protein sim. GBank gij202674 (emb CAB06305 - (Z55120) rHE [Mycobacterium tuberculosis]	Contains protein domain (PF00270) - DEAD/DEAH box helicase	UNCLASSIFIED	35696052, 264769, 264636
287	12745521 (573, 574)			UNCLASSIFIED	284689
288	20756502 (575, 576)	Novel Protein sim. GBank gij765323 (bbs1157676 - (S74439) silk fibroin heavy chain (C-terminal) [Bombyx mori]; silkworms, Peptide Partial, 633 aa) [Bombyx mori]		UNCLASSIFIED	284557
289	80043804 (577, 578)	Novel Protein sim. GBank gij1870009 (emb CAB06680 - (Z92539) hypothetical protein RV1019 [Mycobacterium tuberculosis])	Contains protein domain (PF00440) - Bacterial regulatory proteins, telR family	UNCLASSIFIED	264593, 264600
290	80430175 (579, 580)			UNCLASSIFIED	264766
291	20747431 (581, 582)	Novel Protein sim. GBank gij2508684 (sp P40120 YDCG_ECOLI - 59.4 PROTEIN IN TRG-RIML INTERGENIC REGION PRECURSOR		UNCLASSIFIED	284601
292	80052555 (583, 584)	Novel Protein sim. GBank gij625182 (L39015) - mitochondrial glutamyl-tRNA synthetase [Saccharomyces cerevisiae]		UNCLASSIFIED	264605
293	80082519 (585, 586)	Novel Protein sim. GBank gij1716065 (sp P53528 UVRD_MYCLE - PUTATIVE DNA HELICASE II HOMOLOG	helicase		284908, 284605, 284667, 284689, 284692
294	79830303 (587, 588)	Novel Protein sim. GBank gij117422 (sp P10040 CRB_DROME - CRUMBS PROTEIN PRECURSOR (95F)	Contains protein domain (PF00008) - EGF-like domain	oncogene	35696052, 284806, 265011, 284626, 55611576
295	79444160 (589, 590)	Novel Protein sim. GBank gij1181619 (b BAA11565 - (D82384) a variant of TSC-22 [Gallus gallus]			52844507, 29331622, 284592, 265020, 284639
296	79807076 (591, 592)	Novel Protein sim. GBank gij3649789 (b BAA33403 - (AB012226) SecA [Vibrio elgginolyticus]		synthase	284506
297	79631207 (593, 594)	Novel Protein sim. GBank gij5689867 (emb CAB52004.1 - (AL109863) putative membrane protein [Streptomyces coelicolor A3(2)])		UNCLASSIFIED	284905, 284667, 284638
298	80418698 (595, 596)			UNCLASSIFIED	284905, 284691, 284639, 284766

299	95203298 (597, 598)	Novel Protein sim. GBank gl 220637 dbj BAA01471 - (D10627) zinc finger protein [Mus musculus]	Contains protein domain (PF00086) - Zinc finger, C2H2 type	264486, 263894, 56994075, 22278997, 22278998, 22278999, 20281089, 29331824, 29331825, 29331826, 60432289, 29331827, 29331828, 264905, 264906, 264907, 264908, 52644045, 264909, 264511, 265008, 264910, 264595, 264598, 264758, 33857084, 87188559, 265018, 265019, 264784, 264288, 264768, 264687, 58181582, 264789, 21908765, 21908768, 21908769, 33657023, 264892, 33657109, 27488281, 18106370, 264628, 264829, 55811576, 35695855, 264831, 264634, 264835, 264638, 264639, 63373044, 18108387, 87168516, 22279000, 22279002, 264565, 264588, 264587
300	20711340 (598, 600)	Novel Protein sim. GBank gl 145922 (M20981) - iron dicitrate transport protein precursor [Escherichia coli]	UNCLASSIFIED	264802
301	13511332 (601, 802)	Novel Protein sim. GBank gl 117488 sp P44594 TGT_HAEIN - QUEUINE TRNA-RIBOSYL TRANSFERASE (TRNA-GUANINE TRANSGLYCOSYLASE) (GUANINE INSERTION ENZYME)	transport	264687
302	9875260 (603, 604)	Novel Protein sim. GBank gl 117488 sp P44594 TGT_HAEIN - QUEUINE TRNA-RIBOSYL TRANSFERASE (TRNA-GUANINE TRANSGLYCOSYLASE) (GUANINE INSERTION ENZYME)		264908
303	79574885 (805, 606)	Novel Protein sim. GBank gl 67989 p HJUNVAV - helicase (EC 3.8.1.-) - Autographa californica nuclear polyhedrosis virus	helicase	264689
304	20711344 (607, 608)	Novel Protein sim. GBank gl 728867 sp P40602 APG_ARATH - ANTER-SPECIFIC PROLINE-RICH PROTEIN APG PRECURSOR		264602
305	80412520 (809, 610)	Novel Protein sim. GBank gl 1657554 p PAB16082.1 - (U73857) hypothetical protein [Escherichia coli]		264783
306	8515878 (611, 612)	Novel Protein sim. GBank gl 1657554 p PAB16082.1 - (U73857) hypothetical protein [Escherichia coli]	UNCLASSIFIED	263978
307	80222801 (613, 814)	Novel Protein sim. GBank gl 1710612 sp Q10793 RNH2_MYCTU - PROBABLE RIBONUCLEASE HII (RNASE HII)	UNCLASSIFIED	265010, 21608768, 265020, 18108374, 283977
308	80064305 (615, 816)	Novel Protein sim. GBank gl 1710612 sp Q10793 RNH2_MYCTU - PROBABLE RIBONUCLEASE HII (RNASE HII)	nuclease	264910, 264600, 264605, 264667, 264689, 264638, 16108387
309	80504138 (617, 818)	Novel Protein sim. GBank gl 5420387 emb CAB46678.1 - (AJ243459) proleophosphoglycan [Leishmania major]		264769
310	80053618 (619, 620)	Novel Protein sim. GBank gl 1144522 (U34957) - phosphoribosylaminoimidazole succinocarboxamide synthase [Mycobacterium tuberculosis]		264603
311	11090659 (821, 822)	Novel Protein sim. GBank gl 1144522 (U34957) - phosphoribosylaminoimidazole succinocarboxamide synthase [Mycobacterium tuberculosis]	synthase	264602
312	80054347 (623, 624)	Novel Protein sim. GBank gl 1144522 (U34957) - phosphoribosylaminoimidazole succinocarboxamide synthase [Mycobacterium tuberculosis]	UNCLASSIFIED	264566
313	80048186 (625, 826)	Novel Protein sim. GBank gl 1144522 (U34957) - phosphoribosylaminoimidazole succinocarboxamide synthase [Mycobacterium tuberculosis]		264603, 264587

314	67845112 (627, 628)	Novel Protein sim. GBank gij3661583 (AF092175) - ikaros [Dantio ref]	Contains protein domain (PF00320) - GATA zinc finger	dna_ma_bind	284259, 6043228, 29331828, 264905, 284906, 284908, 284909, 265008, 284910, 6043228, 33657402, 60433438, 33109954, 285011, 265017, 284603, 265018, 284286, 284786, 284892, 35695783, 284828, 284829, 284639, 60170394, 22279002, 284568
315	82356091 (629, 630)	Novel Protein sim. GBank gij1652620(dijBAA17540) - (D90807) pyridine nucleotide transhydrogenase beta subunit [Synecocystis sp.]			284508, 264600, 264782, 264687, 284768, 52844228, 284788, 284688, 284635, 284638, 284638, 284488
316	78911071 (631, 632)	Novel Protein sim. GBank		UNCLASSIFIED	284693
317	20468944 (633, 634)	Novel Protein sim. GBank gij118244(ispP24178)DAPE, ECOLI - SUCCINYL- DIAMINOPIMELATE DESUCCINYLASE (SDAP)		UNCLASSIFIED	284605
318	94141838 (635, 636)	Novel Protein sim. GBank gij4680229(gb)AAD27583.1(AF11827) DNB-5 [Homo sapiens]	Contains protein domain (PF00526) - Dictyostellum (slime mold) repeats	transport	284908, 284909, 284910, 264593, 284594, 284780, 284288, 284788, 284789, 21908769, 284891, 284893, 284828, 65274781, 284835, 284638, 284638, 83373044, 22279002, 284568
319	17289360 (637, 638)	Novel Protein sim. GBank gij1149693(emb)CAA60220] - (X86499) rbsC [Clostridium perfringens]		transport	265018
320	13527675 (639, 640)	Novel Protein sim. GBank gij2611033(isp)005314(GLGC_MYCTU - GLUCOSE-1- PHOSPHATE ADENYLYLTRANSFERASE (ADP- GLUCOSE SYNTHASE) (ADP-GLUCOSE PYROPHOSPHORYLASE)		synthase	264687
321	94134387 (641, 642)	Novel Protein sim. GBank gij1680716 (U68234) - all-trans- retinoic acid 4-hydroxylase [Danio rerio]		cyto450	264508, 264906, 284907, 264908, 265009, 264598, 284784, 284628, 284634, 284635, 284638, 284639, 83373044, 284567
322	68469053 (643, 644)	Novel Protein sim. GBank gij1160355 (U33058) - UNC-89 [Caenorhabditis elegans]		UNCLASSIFIED	55611150, 264681, 60431528, 55810764
323	94553725 (645, 646)			UNCLASSIFIED	284486, 285008, 284593, 284628, 284635
324	78174383 (647, 648)			UNCLASSIFIED	284687
325	79862691 (649, 650)			UNCLASSIFIED	284693
326	28774974 (651, 652)			UNCLASSIFIED	264288, 16106385
327	78778267 (653, 654)	Novel Protein sim. GBank gij451544 (U04267) - proline-rich cell wall protein [Gossypium barbadense]			284488, 284905, 264509, 284910
328	80253202 (655, 656)				
329	10173821 (657, 658)			UNCLASSIFIED	284592
330	88597787 (659, 660)	Novel Protein sim. GBank gij4191358 (AF087825) - claudin- 7 [Mus musculus]		UNCLASSIFIED	284510
331	79754886 (661, 662)	Novel Protein sim. GBank gij80741(pir)S20912 - regulatory protein wh1B - Streptomyces coelicolor		UNCLASSIFIED	284259, 284908
332	60071440 (663, 664)	Novel Protein sim. GBank gij114049(isp)P18480(JAHPF - ALKYL HYDROPEROXIDE REDUCTASE SUBUNIT F (ALKYL HYDROPEROXIDE REDUCTASE F52A PROTEIN)		transcriptfactor	284910, 284687, 284889, 284638, 284567
				reductase	35698423, 284636, 284638, 284565
333	13009555 (665, 666)				284887

334	80230771 (867, 688)	Novel Protein sim. GBank glj322222p[trjS32227 - glutamate dehydrogenase (NADP+)] (EC 1.4.1.4) - <i>Corynebacterium glutamicum</i>	Contains protein domain (PF00208) - Glutamate/Leucine/Phenylalanine/Valine dehydrogenase	dehydrogenase	264905, 264600, 264604, 264488
335	80057028 (889, 670)	Novel Protein sim. GBank glj2183938[embjCA809802] - (Z98800) glpQ2 [<i>Mycobacterium tuberculosis</i>]		esterase	264607, 264603, 264893, 18108374, 264636, 18108387
336	80414319 (671, 672)			UNCLASSIFIED	265009, 264768, 264686
337	11090829 (673, 874)				264602
338	95413134 (875, 676)	Novel Protein sim. GBank glj5454074[re]NP_008303.1pSMRT - silencing mediator for retinoid and thyroid hormone receptors	Contains protein domain (PF00249) - Myb-like DNA-binding domain	nuc1_recpt	264589, 18108397, 22278988, 29331822, 20281099, 29331824, 56182181, 66714117, 29331825, 35896052, 29331828, 264508, 264509, 264905, 264806, 264607, 264908, 264909, 285008, 265008, 264910, 285009, 264758, 55812036, 65274444, 285011, 87188598, 285017, 285018, 265018, 264760, 55811150, 264681, 264762, 18108351, 264682, 264764, 264768, 264885, 264888, 264768, 52644228, 264688, 55811957, 35895917, 264692, 264693, 264628, 18108370, 18108374, 55811578, 35896423, 35895855, 264635, 264555, 264638, 264556, 264637, 264557, 18108380, 264638, 264558, 264639, 18108381, 83373044, 18108385, 87168518, 60432113
339	11388513 (677, 878)	Novel Protein sim. GBank glj4001713[dbj]BAA35087.1 - (AB015879) DnaK [<i>Pseudomonas gingivalis</i>]		eph	264593
340	80504149 (678, 680)	Novel Protein sim. GBank glj2842699[sp]Q92353[UBPC_SCHPO - PUTATIVE UBIQUITIN CARBOXYL-TERMINAL HYDROLASE C6G9.08 (UBIQUITIN THIOLESTERASE) (UBIQUITIN-SPECIFIC PROCESSING PROTEASE) (DEUBIQUITINATING ENZYME)]		ubiquitin	284905, 285019, 264788, 16108374
341	11075198 (881, 682)	Novel Protein sim. GBank glj2688580 (AE001166) - conserved hypothetical protein [Borrelia burgdorferi]	Contains protein domain (PF00290) - Tryptophan synthase alpha chain	isomerase	264605
342	80054196 (683, 684)	Novel Protein sim. GBank glj1864739[embjCAA70601] - (Y09452) YedJ hypothetical protein [<i>Pseudomonas syringae</i>]			264603, 264604
343	20466782 (685, 886)				264605
344	80428870 (887, 688)	Novel Protein sim. GBank glj2117275[embj]CAB09104 - (Z95818) hypothetical protein Rv0807 [<i>Mycobacterium tuberculosis</i>]		UNCLASSIFIED	264600, 264605, 264788, 18108370, 18108374, 35895855
345	80256853 (689, 690)	Novel Protein sim. GBank glj3023317[sp]Q48935[APHA_MYCRA - ACETYL POLYAMINE AMINOHYDROLASE]		histone	264593
346	79831058 (691, 692)	Novel Protein sim. GBank glj4239787[embj]CAA75437 - (Y15166) NADP-glutamate dehydrogenase [<i>Pseudomonas aeruginosa</i>]	Contains protein domain (PF00208) - Glutamate/Leucine/Phenylalanine/Valine dehydrogenase	dehydrogenase	264905

347	79158195 (693, 694)	Novel Protein sim. GBank g 731875 sp P38795 YH4_YEAST - HYPOTHETICAL 80.7 KD PROTEIN IN ERG7-NMD2 INTERGENIC REGION		UNCLASSIFIED	285006, 265008, 265010, 265018, 263987, 283981
348	60020208 (695, 696)	Novel Protein sim. GBank g 1073610 p S47672 - ugpB protein - Escherichia coli		transport	284602, 18108351, 18108387
349	17282112 (697, 698)				265007
350	60502370 (699, 700)	Novel Protein sim. GBank g 3261599 emb CAB009171 - (277137) hypothetical protein Rv1277 [Mycobacterium tuberculosis]		nuclease	285009, 264789, 264869, 16106370
351	80501605 (701, 702)	Novel Protein sim. GBank g 29589367 emb CAA17921 - (AL022117) hypothetical protein [Schizosaccharomyces pombe]		glycoprotein	264789, 264905, 264908
352	11611585 (703, 704)	Novel Protein sim. GBank g 4418302 gb AAD203071 - (AF105716) copia-type pol polyprotein [Zea mays]		protease	264595
353	80061653 (705, 706)	Novel Protein sim. GBank g 1174687 sp P42673 URE1_STAXY - UREASE ALPHA SUBUNIT [UREA AMIDOHYDROLASE]	Contains protein domain (PF00449) - Urease	UNCLASSIFIED	264604
354	56828130 (707, 708)			UNCLASSIFIED	264628
355	80046344 (709, 710)		Contains protein domain (PF00072) - Response regulator receiver domain	UNCLASSIFIED	264909, 264595, 264683, 22279002
356	80043835 (711, 712)			transcript factor	264909, 264591, 264592
357	60070566 (713, 714)	Novel Protein sim. GBank g 497637 (J03939) - cytochrome oxidase d subunit I [Escherichia coli]		oxidase	264605
358	37032756 (715, 716)	Novel Protein sim. GBank g 2280990 (AF006000) - Brg1 [Bordetella pertussis]		UNCLASSIFIED	264768
359	80501488 (717, 718)			UNCLASSIFIED	264604, 264769
360	60026748 (719, 720)			UNCLASSIFIED	264594
361	80584075 (721, 722)	Novel Protein sim. GBank g 3510639 (AF049344) - UDP- GalNAc:polypeptide N-acetylglucosaminyltransferase T5 [Rattus norvegicus]		transferase	22278996, 264259, 29331822, 29331824, 264605, 55811957, 265022
362	13089485 (723, 724)	Novel Protein sim. GBank g 113764 sp P25718 AMY1_ECOLI - ALPHA-AMYLASE PRECURSOR (1,4-ALPHA-D-GLUCAN GLUCANOHYDROLASE)		amylase	264686
363	79750145 (725, 726)				
364	82443593 (727, 728)	Novel Protein sim. GBank g 2829618 sp P95171 NUOK_MYCTU - NADH DEHYDROGENASE I CHAIN K (NADH-UBIQUINONE OXIDOREDUCTASE CHAIN 11) (NUO11)	Contains protein domain (PF00420) - NADH-ubiquinone/plastoquinone oxidoreductase chain 4L	dehydrogenase	264568 264789, 264602, 264604, 264508, 264762, 264638, 264488

365	88040288 (729, 730)	Novel Protein sim. GBank gll4929268[gb AAD33824.1] - (AF144237) LOMP protein [Homo sapiens]	Contains protein domain (PF00412) - LIM domain containing proteins	264486, 21908768, 21908767, 55811576, 21908769, 29148629, 22278995, 22278998, 265020, 265022, 264634, 264691, 264593, 33657023, 33857402, 264693, 264639, 264594, 29331824, 264758, 18108385, 29331827, 87186559, 265018, 22279000, 285019, 264482, 264781, 264681, 18108351, 285017, 264757
366	81821836 (731, 732)	Novel Protein sim. GBank gll4503843[re NP_003908.1]pG2AD - UNKNOWN	Contains protein domain (PF01602) - glycoprotein	60424179, 65274572, 56182575, 22278994, 58994075, 22278998, 264259, 29331822, 29331824, 56182181, 80424288, 68714117, 29331825, 80432289, 29331826, 29331827, 29331828, 264905, 264628, 56182435, 265008, 284512, 285008, 264591, 55812038, 55811386, 285010, 87186559, 265017, 265018, 284804, 265019, 55811150, 264448, 264389, 264288, 264888, 264768, 56181582, 21908788, 21908789, 55811857, 35895617, 265022, 60170815, 33657023, 65274620, 16108365, 263987, 33657109, 33657349, 35895783, 264628, 18108378, 55811576, 65274791, 35895855, 56182323, 83373044, 60432113, 264563, 264584, 264587, 264509
368	79607265 (735, 736)	Novel Protein sim. GBank gll3913029[sp P94967 ALR_MYCSM - ALANINE RACEMASE	UNCLASSIFIED	264508, 264604, 264805, 264636
369	95292917 (737, 738)	Novel Protein sim. GBank gll3249549 (AF018261) - EH domain binding protein Epsin [Rattus norvegicus]		264905, 264582, 264605, 264766, 264691
370	68090865 (739, 740)	Novel Protein sim. GBank gll2995299[emb CAA18328] - (AL022268) putative tRNA delta(2)-isopentenylpyrophosphate transferase [Streptomyces coelicolor]	Contains protein domain (PF01715) - transferase IPP transferase	264905, 264906, 264510, 264600, 264601, 264602, 264603, 265018, 264604, 264605, 265021, 264892, 264836, 264584
371	95292599 (741, 742)	Novel Protein sim. GBank gll1805409[db BAA08970] - (D50453) homologues to nitrate hydratase region 3'-hypothetical protein P47K of P. chlororaphis [Bacillus subtilis]		264584
372	80021107 (743, 744)	Novel Protein sim. GBank gll2508393[sp P31576 FIXX_ECOLI - FERREDOXIN LIKE PROTEIN	UNCLASSIFIED	264909
373	79863768 (745, 746)	Novel Protein sim. GBank gll3341640[emb CAA13164] - (AJ231122) z811 [Vibrio cholerae]	UNCLASSIFIED	264905, 264906
374	79847568 (747, 748)	Novel Protein sim. GBank gll5456934[gb AAD43716.1] - (AF152322) protocadherin gamma A2 [Homo sapiens]	cadherin	65274572, 264259, 29331826, 56182435, 60433356, 80433438, 264757, 55812038, 264758, 55811957, 264690, 33857023
375	91230181 (749, 750)	Novel Protein sim. GBank gll1805409[db BAA08970] - (D50453) homologues to nitrate hydratase region 3'-hypothetical protein P47K of P. chlororaphis [Bacillus subtilis]	UNCLASSIFIED	264768
376	80505214 (751, 752)			
377	10339063 (753, 754)			264906

378	60056153 (755, 756)	Novel Protein sim. GBank g 1076013 p j A49930 - carB protein homolog - Mycobacterium bovis (strain BCG) (fragment)	Contains protein domain (PF00289) - Carbamoyl-phosphate synthase (CPSase)	UNCLASSIFIED	265006, 264555
379	80503437 (757, 758)	Novel Protein sim. GBank g 218556 d b j BA021741 - (D12651) glucose dehydrogenase [Escherichia coli]	Contains protein domain (PF01011) - PQQ enzyme repeat	dehydrogenase	264789
380	80060937 (759, 760)	Novel Protein sim. GBank g 3327136 d b j BA031636i - (AB014581) KIAA0681 protein [Homo sapiens]		UNCLASSIFIED	264604
381	11769027 (761, 762)				264684
382	80054377 (763, 764)				264592
383	83259025 (765, 766)				264595, 265017, 265021, 264638, 87188518, 22279002
384	95314255 (767, 768)			UNCLASSIFIED	264286, 264786, 263987, 85274791, 35695855, 263981, 83373044, 264567
385	10237679 (769, 770)				264692
386	76833434 (771, 772)	Novel Protein sim. GBank g 1073456 p j S47810 - probable alcohol dehydrogenase (EC 1.1.1) - Escherichia coli	Contains protein domain (PF00465) - Iron-containing alcohol dehydrogenases	dehydrogenase	264908
387	17960637 (773, 774)	Novel Protein sim. GBank g 1460074 emb CA010491 - (Z77250) hypothetical protein Rv2566 [Mycobacterium tuberculosis]	Contains protein domain (PF01841) - Transglutaminase-like superfamily	UNCLASSIFIED	264760
388	67741376 (775, 776)	Novel Protein sim. GBank g 4240169 d b j BA074863.11 - (AB020647) KIAA0640 protein [Homo sapiens]	Contains protein domain (PF00846) - F-box domain.	homeobox	35698286, 264905, 66712502, 60432229, 284593, 60433358, 264888, 264888, 21908765, 264891, 22279000, 264462
389	79316971 (777, 778)			UNCLASSIFIED	18108394, 22278998, 264630, 264556, 22278002
390	80079949 (779, 780)			UNCLASSIFIED	264600
391	7657302 (781, 782)	Novel Protein sim. GBank g 654065 emb CAA563371 - (X63413) U88 [Human herpesvirus 6]			264482
392	79796056 (783, 784)			UNCLASSIFIED	264908
393	33206031 (785, 786)	Novel Protein sim. GBank g 3376523 emb CAA088671 - (AJ009632) cyclomaltoextrinase glucanotransferase [Thermotoga neopolitana]	Contains protein domain (PF00047) - synthase	synthase	264802, 21906764
394	10104463 (787, 788)				264893
395	60228010 (789, 790)			UNCLASSIFIED	264508, 264563
396	20436224 (791, 792)	Novel Protein sim. GBank g 2677780 (U70327) - unknown [Pareitropus polyacis]	Contains protein domain (PF00047) - stitut	UNCLASSIFIED	264558
397	80417014 (793, 794)	Novel Protein sim. GBank g 4507809 ref NP_000368.1 pWAS - Wiskott-Aldrich syndrome (eczema-thrombocytopenia)	Immunoglobulin domain		265007, 265009, 264508, 264556, 264629, 264786
398	91230517 (795, 796)	Novel Protein sim. GBank g 1518456 (U45998) - mitochondrial solute carrier [Onchocerca volvulus]	Contains protein domain (PF00153) - Mitochondrial carrier proteins	transport	18108398, 22278995, 22278996, 56994075, 22278999, 264259, 26331824, 26331828, 264905, 264908, 265007, 265008, 265009, 21906754, 33657084, 265017, 264446, 264286, 264788, 21906765, 21906768, 21906767, 265020, 265021, 33637023, 33657109, 264628, 35698423, 35695855, 264952, 18108380, 264567, 18108391

398	80055278 (797, 798)	Novel Protein sim. GBank gij3356091[dbj BA31995] - (AB015974) glycerol kinase [Pseudomonas tolaasii]	Contains protein domain (PF00370) - kinase FGGY family of carbohydrate kinases	284592, 284595
400	94117480 (799, 800)	Novel Protein sim. GBank gij728635[sp P39192 ALLJ5_HUMAN - IIII ALU SUBFAMILY SC WARNING ENTRY IIII	Contains protein domain (PF00580) - cadherin Leucine Rich Repeat	18108394, 56182575, 22278995, 22278997, 22278999, 284259, 29331824, 265008, 265007, 265009, 60432229, 33657402, 21908754, 285010, 265017, 265018, 265019, 18108351, 18108357, 21908765, 265021, 285022, 284891, 284892, 33657023, 18108370, 65274791, 284634, 284836, 60170394, 58182323, 284594
401	11397491 (801, 802)	Novel Protein sim. GBank gij4928292[gb AA033527.1 AF13211] FhuA [Staphylococcus aureus]	transport	
402	95420284 (803, 804)	Novel Protein sim. GBank gij5689487[dbj BA383027.1] - (AB028998) KIAA1075 protein [Homo sapiens]	Contains protein domain (PF00017) - phosphatase Src homology domain 2	65274572, 56182575, 35696286, 22278998, 22278999, 284093, 284259, 29331822, 29331824, 29331825, 29331826, 60432289, 29331827, 29331828, 284908, 284907, 284909, 285008, 284511, 285007, 285008, 284910, 284591, 33657402, 80433356, 60433438, 284596, 21906754, 52844296, 265010, 265011, 87168559, 265017, 285018, 265019, 284881, 18108351, 284682, 284448, 284288, 284684, 284766, 284767, 284686, 21908765, 21908766, 21908767, 21908768, 21908769, 55811957, 265020, 265021, 285022, 284690, 284693, 85274620, 35695763, 284628, 18108370, 284829, 18108379, 35698423, 55811576, 284835, 284638, 284557, 284839, 18108385, 22279002, 284563, 284564, 284585, 284586, 284768, 284632, 284639, 284563, 284682, 285009, 284882
403	80439913 (805, 806)		UNCLASSIFIED	
404	11809865 (807, 808)		polymerase	
405	79471280 (809, 810)	Novel Protein sim. GBank gij266164[emb CAA15755] - (AL009198) dnaE2 [Mycobacterium tuberculosis]	Contains protein domain (PF00159) - Pancreatic hormone peptidases	18108357, 284693
406	79634172 (811, 812)		UNCLASSIFIED	
407	80478229 (813, 814)		UNCLASSIFIED	284769
408	80078958 (815, 816)		UNCLASSIFIED	284600
409	5640527 (817, 818)	Novel Protein sim. GBank gij3047117 (AF058819) - similar to ATP-dependent RNA helicases [Arabidopsis thaliana]	helicase	284259

410	95357486 (619, 620)	Novel Protein sim. GBank gij475016[idj]BAA08164] - (D28801) Unknown [Mus musculus]	UNCLASSIFIED	264489, 52646365, 52846842, 56161886, 35696286, 52845080, 29331822, 29331824, 56162161, 29331625, 60424269, 35696052, 33658970, 264506, 264509, 264905, 264906, 264907, 264908, 52644045, 264909, 264510, 265007, 264512, 265008, 264910, 33657402, 264758, 52646317, 55811366, 265010, 265011, 265017, 264604, 265018, 55811150, 264782, 264784, 264786, 264687, 264788, 264769, 52844229, 21906766, 265020, 265021, 264534, 52844150, 264892, 33657023, 65274620, 33657109, 33657182, 27486261, 35695763, 264626, 264628, 60431528, 18108376, 263978, 35696423, 35895655, 264632, 264634, 264635, 264637, 264638, 264558, 264639, 56162323, 264559, 60432113, 22278002, 264563, 264565, 264486
411	80501670 (621, 622)		UNCLASSIFIED	264769
412	80241662 (623, 624)	Novel Protein sim. GBank gij3261764[emb]CAB08997] - (Z95558) htpX [Mycobacterium tuberculosis]	eph	264907, 264910, 263973, 22278002
413	11076446 (625, 626)	Novel Protein sim. GBank		264605
414	82050564 (627, 628)	gij129038[sp]P20707[ODO1_AZOV] - 2-OXOGLUTARATE DEHYDROGENASE E1 COMPONENT (ALPHA-KETOGLUTARATE DEHYDROGENASE)	dehydrogenase	18108374, 264760, 264769, 264602, 264638, 264603, 264909, 264805
415	84453144 (629, 630)	Novel Protein sim. GBank gij4668350[gb]AAD31273.1[AF132025] rhophilin [Drosophila melanogaster]	UNCLASSIFIED	264806, 87188516
416	80402775 (631, 632)	Novel Protein sim. GBank gij2555172 (AF025543) - ArcC; carbamate kinase [Rhizobium etli]	kinase	264498, 264600, 264602, 264764, 264638
417	20153797 (633, 634)	Novel Protein sim. GBank gij1709171[sp]P52311[MTX2_XANOR - MODIFICATION METHYLASE XORII (CYTOSINE-SPECIFIC METHYLTRANSFERASE XORII) (M.XORII)]		264805
418	94125641 (635, 636)		UNCLASSIFIED	
419	95314273 (637, 638)	Novel Protein sim. GBank gij3261659[emb]CAB03751] - (Z61368) hypothetical protein Rv2419c [Mycobacterium tuberculosis]	collagen	264689, 264693
420	37036349 (639, 640)		phosphatase	264908, 264910, 264764, 264639
421	95292942 (641, 642)	Novel Protein sim. GBank gij2916942[emb]CAA17580] - (AL021999) hypothetical protein Rv0961 [Mycobacterium tuberculosis]	phosphatase	264908, 264800, 264801, 264803, 264604, 264760, 264769
422	79471283 (643, 644)	Novel Protein sim. GBank gij231752[sp]Q00767[CH61_STRAL - 80 KD CHAPERONIN 1 (PROTEIN CPN60.1) (GROEL PROTEIN 1) (HSP59)]	eph	22278096, 264662, 18108376, 18108367
423	79604948 (645, 646)		UNCLASSIFIED	264509

424	78988557 (847, 848)	Novel Protein sim. GBank g 4826814 ref NP_004977.1 p KTN1 - kinesin receptor		strud	265019
425	80431450 (849, 850)	Novel Protein sim. GBank g 1703701 p bs 178462 - KRP5-kinesin-related protein [rats, testes, Peptide Partial, 167 aa]	Contains protein domain (PF00225) Kinesin motor domain	strud	264808, 265007, 55611386, 264768, 55610764
426	80064522 (851, 852)	Novel Protein sim. GBank		UNCLASSIFIED	264605, 264559
427	80057232 (853, 854)	Novel Protein sim. GBank g 231629 p P29929 COBN_PSEDE - COBN PROTEIN		UNCLASSIFIED	264803, 264636
428	79467798 (855, 856)	Novel Protein sim. GBank g 181266 p j S22697 - extensin - Volvox carter (fragment)		UNCLASSIFIED	264683
429	80091252 (857, 858)	Novel Protein sim. GBank g 1808154 p mb CA806451 - (294395) hypothetical protein Rv0688 [Mycobacterium tuberculosis]		UNCLASSIFIED	35698423, 35695763, 35695655, 265017, 264584, 264762
430	80504192 (859, 860)	Novel Protein sim. GBank g 1808154 p mb CA806451 - (294395) hypothetical protein Rv0688 [Mycobacterium tuberculosis]		reductase	264508, 264805, 264509, 264908, 264909, 265008, 264800, 264687, 264769, 264689, 264636, 264638, 18106365, 264488
431	20524249 (861, 862)	Novel Protein sim. GBank g 3123552 p mb CAA18608 - (AL022578) dJ393P12.2 (hypothetical Proline-rich protein KIAA0269 LIKE) [Homo sapiens]		UNCLASSIFIED	264566
432	18525372 (863, 864)	Novel Protein sim. GBank		UNCLASSIFIED	265020
433	61484303 (865, 866)	Novel Protein sim. GBank g 2495272 p Q99826 CDX2_HUMAN - HOMEBOX PROTEIN CDX-2 (CAUDAL-TYPE HOMEBOX PROTEIN 2) [CDX-3]	Contains protein domain (PF00189) PH domain	UNCLASSIFIED	264907, 264908, 264909, 264910, 264592, 264595, 264758, 264604, 264762, 264763, 264636, 264637, 22278002
434	94326323 (867, 868)	Novel Protein sim. GBank g 114105 p P09532 ARAH_ECOLI - L-ARABINOSE TRANSPORT SYSTEM PERMEASE PROTEIN ARAH		UNCLASSIFIED	55812038, 56182161, 56161562, 29331826, 35696052, 55610764, 55611576, 65274791, 35695855, 80432113, 55611150, 264636, 264766
435	80502738 (869, 870)	Novel Protein sim. GBank		transport	264595, 264769
436	41085953 (871, 872)	Novel Protein sim. GBank		UNCLASSIFIED	265020, 22279002
437	11399281 (873, 874)	Novel Protein sim. GBank		UNCLASSIFIED	264593
438	11773835 (875, 876)	Novel Protein sim. GBank g 3242702 (AC003040) - hypothetical protein [Arabidopsis thaliana]		UNCLASSIFIED	264686
439	80019495 (877, 878)	Novel Protein sim. GBank g 3242702 (AC003040) - hypothetical protein [Arabidopsis thaliana]		UNCLASSIFIED	264905, 264800, 264602, 264604
440	79841082 (879, 880)	Novel Protein sim. GBank g 2291232 p gb A885351.1 - (AF016427) Contains similarity to Pfam domain: PF00004 (AAA). Score=266.1, E-value=3.7e-77, N=1 [Caenorhabditis elegans]	Contains protein domain (PF00004) - ATPases associated with various cellular activities (AAA)	ATPase-associated	35698052, 264905, 264908, 264909, 265011, 35698423
441	20386935 (881, 882)	Novel Protein sim. GBank g 5639946 p b A045904.1 AF16132 - (AF161328) histidine kinase CstS [Corynebacterium diptheriae]			264805
442	65281056 (883, 884)	Novel Protein sim. GBank g 1184790 (U46088) - von Ebner minor salivary gland protein [Mus musculus]		UNCLASSIFIED	29331830, 264909
443	82456427 (885, 886)	Novel Protein sim. GBank g 5669933 p mb CA852056.1 - (AL109732) putative ATP-binding RNA helicase [Streptomyces coelicolor A3(2)]		UNCLASSIFIED	35696052, 264508, 264906, 264512, 264604, 264762, 264769, 264689, 264636
444	11395897 (887, 888)	Novel Protein sim. GBank g 1763249 p b BAA11726 - (B03026) homologous to citrate-sodium symport (citrate transporters); hypothetical [Bacillus subtilis]		UNCLASSIFIED	264591

445	79552708 (889, 890)	Novel Protein sim. GBank gij5531272[emb]CAB50897.1] - (AJ243800) WSC4 homologue [Kluyveromyces fragilis]		UNCLASSIFIED	264693
446	79810937 (891, 892)	Novel Protein sim. GBank gij538413 (L36315) - zinc finger protein [Mus musculus]	Contains protein domain (PF00086) - Zinc finger, C2H2 type		264508
447	80438888 (893, 894)	Novel Protein sim. GBank gij542814[emb]CAB02185] - (Z80108) fnt [Mycobacterium tuberculosis]	Contains protein domain (PF00551) - Formyl transferase	transcript factor	264768, 55811578
448	80238110 (895, 896)	Novel Protein sim. GBank gij118764[sp]P104-43[DP3A_ECOLI - DNA POLYMERASE III, ALPHA CHAIN		dehydrogenase	264508, 264600, 264603, 264605, 264682, 264769, 18108362, 264634, 18108387
449	20460634 (897, 898)	Novel Protein sim. GBank gij4589508[db]BAA76775.1] - (AB023148) KIAA0931 protein [Homo sapiens]		polymerase	264805, 264559
450	94631210 (899, 900)	Novel Protein sim. GBank gij2850814 (AE001104) - conserved hypothetical protein [Archaeoglobus fulgidus]	Contains protein domain (PF00481) - Protein phosphatase 2C	phosphatase	65274572, 22278988, 28331824, 28331825, 284906, 264910, 264992, 52848317, 285017, 21808767, 55811957, 56528488, 22279002
451	21435609 (901, 902)	Novel Protein sim. GBank gij2493000[sp]Q08450[SCOT_CAEEL - PROBABLE SUCCINYL-COA:3-KETOACID-COENZYME A TRANSFERASE PRECURSOR (3-OXOACID COA-TRANSFERASE)		UNCLASSIFIED	264486
452	10287278 (903, 904)	Novel Protein sim. GBank gij118764[sp]P104-43[DP3A_ECOLI - DNA POLYMERASE III, ALPHA CHAIN		UNCLASSIFIED	264692
453	52560098 (905, 906)	Novel Protein sim. GBank gij4589508[db]BAA76775.1] - (AB023148) KIAA0931 protein [Homo sapiens]		UNCLASSIFIED	264907, 264600
454	39523922 (907, 908)	Novel Protein sim. GBank gij4589508[db]BAA76775.1] - (AB023148) KIAA0931 protein [Homo sapiens]		transferase	264603
455	13088692 (909, 910)	Novel Protein sim. GBank gij4589508[db]BAA76775.1] - (AB023148) KIAA0931 protein [Homo sapiens]		UNCLASSIFIED	284687
456	79563081 (911, 912)	Novel Protein sim. GBank gij4589508[db]BAA76775.1] - (AB023148) KIAA0931 protein [Homo sapiens]		UNCLASSIFIED	284691
457	79831273 (913, 914)	Novel Protein sim. GBank gij4589508[db]BAA76775.1] - (AB023148) KIAA0931 protein [Homo sapiens]		UNCLASSIFIED	264805
458	79581227 (915, 916)	Novel Protein sim. GBank gij4589508[db]BAA76775.1] - (AB023148) KIAA0931 protein [Homo sapiens]	Contains protein domain (PF00595) - PDZ domain (Also known as DHR or GLGF)	kinase	55812038, 265010, 265018, 264681
459	80567359 (917, 918)	Novel Protein sim. GBank gij4589508[db]BAA76775.1] - (AB023148) KIAA0931 protein [Homo sapiens]	Contains protein domain (PF00130) - Phorbol esters/diacylglycerol binding domain (C1 domain)	kinase	22278987, 264258, 28331828, 285018, 264448, 264369, 21806765, 35686423
460	79245680 (919, 920)	Novel Protein sim. GBank gij113158[sp]P25516[ACO1_ECOLI - ACONITATE HYDRATASE 1 (CITRATE HYDRO-LYASE 1) (ACONITASE 1)]		UNCLASSIFIED	264906
461	95287618 (921, 922)	Novel Protein sim. GBank gij1188574[sp]P42464[ATPB_CORGL - ATP SYNTHASE BETA CHAIN		synthase	264602, 264605, 264768, 264769, 265021, 33657023, 264559

462	79606589 (923, 924)	Novel Protein sim. GBank gl 1348891 sp P45597 PTF1_XANCP - MULTIPHOSPHORYL TRANSFER PROTEIN (MTP) (CONTAINS: PHOSPHOENOLPYRUVATE-PROTEIN PHOSPHOTRANSFERASE (PHOSPHOTRANSFERASE SYSTEM, ENZYME I); PHOSPHOCARRIER PROTEIN HPR (PROTEIN H); PTS SYSTEM, FRUCTOSE-SPECIFIC IIA COMPONENT ...	Contains protein domain (PF00391) - PEP-utilizing enzymes	284807
463	79786417 (925, 926)	Novel Protein sim. GBank gl 854065 emb CAA583371 - (X83413) U88 [Human herpesvirus 8]	UNCLASSIFIED	284905, 284806, 284908, 284909, 284910, 284581, 284595, 285011, 284632, 284635, 284638, 284637, 284638, 284639 284634
464	82340151 (927, 929)	Novel Protein sim. GBank gl 5689778 emb CAB62137.1 - (AJ242832) calpain [Homo sapiens]	UNCLASSIFIED	285017, 21906764, 285020
465	83005730 (929, 930)	Novel Protein sim. GBank gl 1806175 emb CAB06470 - (Z84395) psc [Mycobacterium tuberculosis]	cathepsin	
466	20460645 (931, 932)	Novel Protein sim. GBank gl 1806175 emb CAB06470 - (Z84395) psc [Mycobacterium tuberculosis]	Contains protein domain (PF00417) - Ribosomal protein S3, N-terminal domain.	284605, 284559
467	80409035 (933, 934)	Novel Protein sim. GBank gl 548705 sp P38949 RBSB_BACSU - D-RIBOSE-BINDING PROTEIN PRECURSOR	UNCLASSIFIED	264764
468	52582208 (935, 936)	Novel Protein sim. GBank gl 2114024 emb CAB08957 - (Z85558) gpcC1 [Mycobacterium tuberculosis]	UNCLASSIFIED	284692
469	19520527 (937, 938)	Novel Protein sim. GBank gl 2808459 emb CAA17347 - (AL021929) cobQ [Mycobacterium tuberculosis]	UNCLASSIFIED	284488
470	80502756 (938, 940)	Novel Protein sim. GBank gl 114921 sp P17447 BETT_ECOLI - HIGH-AFFINITY CHOLINE TRANSPORT PROTEIN	synthase	284602, 284769
471	17937351 (941, 942)	Novel Protein sim. GBank gl 114921 sp P17447 BETT_ECOLI - HIGH-AFFINITY CHOLINE TRANSPORT PROTEIN	transport	265019
472	80047458 (943, 944)	Novel Protein sim. GBank gl 862343 (L10808) - Gcap1 gene product [Mus musculus]	UNCLASSIFIED	284596, 284685, 284557 284369
473	20558793 (945, 946)	Novel Protein sim. GBank gl 862343 (L10808) - Gcap1 gene product [Mus musculus]	UNCLASSIFIED	22278997, 284882, 284288
474	80593385 (947, 948)	Novel Protein sim. GBank gl 5453656 ref NP_006329.1 pGAC1 - glioma amplified on chromosome 1 protein (leudne-rich)	UNCLASSIFIED	264807, 284808, 284511, 285009, 284762, 284446, 284638, 284638
475	82454665 (948, 950)	Novel Protein sim. GBank gl 5453656 ref NP_006329.1 pGAC1 - glioma amplified on chromosome 1 protein (leudne-rich)	glycoprotein	65274572, 60432049, 284259, 284508, 52644045, 55812038, 284758, 285011, 284288, 284686, 52844229, 65274761, 284638, 284556
476	94143857 (951, 952)	Novel Protein sim. GBank gl 5453656 ref NP_006329.1 pGAC1 - glioma amplified on chromosome 1 protein (leudne-rich)	UNCLASSIFIED	284638
477	79175833 (953, 954)	Novel Protein sim. GBank gl 1127551 (U16939) - orf2 [Batrachocottus baikalensis]	UNCLASSIFIED	264690, 264693
478	79633483 (955, 956)	Novel Protein sim. GBank gl 1127551 (U16939) - orf2 [Batrachocottus baikalensis]	collagen	284686, 35685855, 285008, 284631, 284910, 284632, 284838, 285018, 284369, 284909
479	80189746 (957, 958)	Novel Protein sim. GBank gl 1127551 (U16939) - orf2 [Batrachocottus baikalensis]	mapolymerase	264369
480	79390729 (959, 960)	Novel Protein sim. GBank gl 1127551 (U16939) - orf2 [Batrachocottus baikalensis]	UNCLASSIFIED	264693
481	79624578 (961, 962)	Novel Protein sim. GBank gl 4083042 (AF068085) - GP000; much-like glycoprotein [Cryptosporidium parvum]	UNCLASSIFIED	284909, 284686, 284768, 284893, 55811576, 56182323, 18108385
482	83050811 (963, 964)	Novel Protein sim. GBank gl 4083042 (AF068085) - GP000; much-like glycoprotein [Cryptosporidium parvum]	UNCLASSIFIED	

483	20283306 (965, 966)	Novel Protein sim. GBank gjl2104303[embjCAB08632] - (Z95387) hypothetical protein Rv2810c [Mycobacterium tuberculosis]	Contains protein domain (PF00534) - Glycosyl transferases group 1	284600
484	11618046 (967, 968)	Novel Protein sim. GBank gjl3450883 (AF083334) - fibroin [Antheraea pernyi]	UNCLASSIFIED	284594
485	80191234 (969, 970)	Novel Protein sim. GBank gjl5042272[embjCAB44526.1] - (AL076818) nuof, NADH dehydrogenase subunit [Streptomyces coelicolor]	UNCLASSIFIED	284369, 21806765, 22279000, 22278002
486	80059042 (971, 972)		dehydrogenase	284604
487	11613339 (973, 974)			284638
488	91222383 (975, 976)	Novel Protein sim. GBank gjl5724778[gbjAAC53522.2] - (AF012273) rho-type GTPase-activating protein rhoGAPX-1 [Mus musculus]	Contains protein domain (PF00620) - RhoGAP domain	284886, 66714117, 284768, 18108385, 55811578, 265006, 265008, 265009, 265019, 22279002, 284259, 18108370, 284907, 284764, 56182323, 284288, 284893
489	10867710 (977, 978)	Novel Protein sim. GBank gjl3882223[dbjBAA34471.1] - (AB018294) KIAA0751 protein [Homo sapiens]	kinase	284639
490	95361124 (979, 980)	Novel Protein sim. GBank gjl82081[pirjA25494 - hydroxyproline-rich glycoprotein - tomato (fragment)]	collagen	22278998, 29331822, 29331828, 284107, 284908, 284110, 285009, 284592, 284593, 60433356, 284288, 284693, 263874, 263978, 20281071, 80432113
491	80486412 (981, 982)	Novel Protein sim. GBank gjl2894206[embjCAA17072] - (AL021840) hypothetical protein Rv3256c [Mycobacterium tuberculosis]	UNCLASSIFIED	284769
492	87421264 (983, 984)			284600
493	11692942 (985, 986)		UNCLASSIFIED	284638
494	87726604 (987, 988)	Novel Protein sim. GBank gjl5262605[embjCAB45743.1] - (AL080150) hypothetical protein [Homo sapiens]	UNCLASSIFIED	284489, 35696286, 60432289, 29331828, 35696052, 284509, 284905, 284906, 284907, 284908, 284909, 284510, 284511, 285009, 284910, 33657402, 284762, 284764, 284768, 284789, 284688, 21908785, 21908769, 35695917, 285020, 284693, 33657109, 284629, 35696423, 35695855, 284634, 284638
495	80028599 (989, 990)	Novel Protein sim. GBank gjl2781517[embjCAA16054] - (AL021246) hypothetical protein Rv2477c [Mycobacterium tuberculosis]	Contains protein domain (PF00005) - ABC transporter	284602, 284862, 284838
496	78985624 (991, 992)	Novel Protein sim. GBank gjl230281[dbjIR69] - 434 Repressor (Amino-Terminal Domain) (R1-69)	Contains protein domain (PF01381) - Helix-turn-helix	284601, 265021
497	78948661 (993, 994)	Novel Protein sim. GBank gjl128738[spP28225]PDXH_ECOLI - PYRIDOXAMINE 5'-PHOSPHATE OXIDASE (PNP/IMP OXIDASE)	oxidase	285006

498	80095488 (895, 998)	Novel Protein sim. GBank gi 1145789 (U41662) - neurotigin 2 [Rattus norvegicus]	Contains protein domain (PF00135) - Carboxylesterases	esterase	264259, 26331828, 35698052, 264506, 264509, 264805, 264806, 264807, 264808, 264909, 264510, 264511, 265009, 264910, 264591, 33857402, 264758, 265010, 265011, 264800, 264601, 264605, 264683, 264764, 264766, 264787, 264788, 264887, 264769, 21906767, 33657023, 264693, 264828, 264629, 35698423, 264630, 264832, 264834, 264835, 264837, 264838, 264558, 264839, 18108385, 264563, 264564, 264565, 264566, 264587, 264605
499	20436222 (997, 998)	Novel Protein sim. GBank gi 97480 pir S19739 - Integral membrane protein - Rhodobacter capsulatus		UNCLASSIFIED	264605
500	11076810 (899, 1000)				264605
501	13418034 (1001, 1002)	Novel Protein sim. GBank gi 5708250 emb CAB52363.1 - (AL108747) putative Integral membrane protein [Streptomyces coelicolor A3(2)]		UNCLASSIFIED	264688
502	80021176 (1003, 1004)	Novel Protein sim. GBank gi 468878 emb CAB38132.1 - (AL035581) glucose-6-phosphate isomerase [Streptomyces coelicolor]	Contains protein domain (PF00342) - Phosphoglucose isomerase	isomerase	22278996, 265011, 264602, 264605, 264635
503	20264483 (1005, 1006)			UNCLASSIFIED	264584
504	10887321 (1007, 1008)			UNCLASSIFIED	264887
505	95003068 (1009, 1010)			UNCLASSIFIED	264369
506	16454292 (1011, 1012)	Novel Protein sim. GBank gi 4033509 sp P02598 CALM_TETPY - CALMODULIN EF hand	Contains protein domain (PF00036) - EF hand	struct	265010
507	20451588 (1013, 1014)	Novel Protein sim. GBank gi 2501069 sp Q46127 SYW_CLOLO - TRYPTOPHANYL-TRNA SYNTHETASE (TRYPTOPHAN-TRNA LIGASE) (TRPRS)		UNCLASSIFIED	264604
508	70841424 (1015, 1016)	Novel Protein sim. GBank gi 468088 sp P34618 YO82_CAEEL - HYPOTHETICAL 33.8 KD PROTEIN ZK1236.2 IN CHROMOSOME III		UNCLASSIFIED	284808
509	11776388 (1017, 1018)				264636
510	83373465 (1019, 1020)			UNCLASSIFIED	264887, 264839
511	16525578 (1021, 1022)				265007
512	20388484 (1023, 1024)	Novel Protein sim. GBank gi 2497419 sp P55635 Y4RB_RHISN - PUTATIVE INTEGRASERECOMBINASE Y4RB		UNCLASSIFIED	264565
513	79457404 (1025, 1026)	Novel Protein sim. GBank gi 1276887 (U41809) - cyclin J [Drosophila melanogaster]	Contains protein domain (PF00134) - Cyclin	cyclin	264683, 264889, 35698423, 264639
514	79813805 (1027, 1028)	Novel Protein sim. GBank gi 1184780 (U46068) - von Ebner minor salivary gland protein [Mus musculus]		UNCLASSIFIED	29331830, 264909
515	79462591 (1029, 1030)				22278998, 264690
516	8662020 (1031, 1032)	Novel Protein sim. GBank gi 2127400 pir S65770 - maltotrioglycosyltrehalose trehalohydrolase - Arthrobacter sp. (strain Q36)		anyase	284910

517	95292984 (1033, 1034)	Novel Protein sim. GBank gij2883605 (AE000725) - ribose 5 phosphate isomerase B [Aquilifex aceticus]				265018, 264605, 264764, 264766, 264687, 264891, 264585
518	8491831 (1035, 1036)	Novel Protein sim. GBank gij854065 [amb]CAA583371 - (X83413) U88 [Human herpesvirus 8]			UNCLASSIFIED	264487
519	91677888 (1037, 1038)	Novel Protein sim. GBank gij5689365 [obj]BAA83073.1 - (AB024075) B120 [Homo sapiens]	Contains protein domain (PF01388) - ARID DNA binding domain	dna_ma_bind		52644507, 22278997, 22278998, 80432049, 264259, 52645080, 29331826, 35896052, 264905, 60424289, 29331826, 35896052, 264905, 29331830, 66712502, 264511, 265007, 264591, 60432229, 33857402, 80433438, 21906754, 33109954, 52844296, 87168474, 87168559, 265017, 265018, 264804, 285019, 264681, 264448, 284389, 264288, 264885, 21906765, 21906786, 21906767, 21906769, 21906765, 21906786, 21906767, 21906769, 265021, 60170615, 33857023, 264692, 52645129, 33857109, 27488282, 27486264, 35695763, 18108370, 264629, 52644332, 56182323, 264639, 63373044, 18108385, 56528486, 80432113
520	79869188 (1039, 1040)	Novel Protein sim. GBank gij1169126 [sp]P46839 [CTPA_MYCLE - CATION-TRANSPORTING P-TYPE ATPASE A]			transport	264768
521	11076821 (1041, 1042)	Novel Protein sim. GBank gij1172869 [sp]P44331 [RBSK_HAEIN - RIBOKINASE]	Contains protein domain (PF00294) - ptkB family carbohydrate kinase	kinase		264605
522	80435060 (1043, 1044)	Novel Protein sim. GBank gij1172869 [sp]P44331 [RBSK_HAEIN - RIBOKINASE]				264905, 264768
523	18356013 (1045, 1046)	Novel Protein sim. GBank gij2132243 [pji]S81028 - hypothetical protein YPL238c - yeast (Saccharomyces cerevisiae)			UNCLASSIFIED	264628
524	80261805 (1047, 1048)	Novel Protein sim. GBank gij4033608 [obj]BAA351361 - (AB012309) B2HC [Anthracidaris cressispina]			ATPase-associated	284082, 264596, 265011
525	79810046 (1049, 1050)	Novel Protein sim. GBank gij4106610 [emb]CAA213651 - (AL031866) ORF42, len=388 aa, similarity to an aminotransferase, in P95987 Sulfolobus solfataricus, (401 aa), 33.1% identity in 393 aa overlap. Fasta scores: opt:468, E0: 8.5e-24, in Q64602 R. norvegicus, (425 aa), 28.8% identi...			UNCLASSIFIED	284907
526	38827630 (1051, 1052)	Novel Protein sim. GBank gij4106610 [emb]CAA213651 - (AL031866) ORF42, len=388 aa, similarity to an aminotransferase, in P95987 Sulfolobus solfataricus, (401 aa), 33.1% identity in 393 aa overlap. Fasta scores: opt:468, E0: 8.5e-24, in Q64602 R. norvegicus, (425 aa), 28.8% identi...			UNCLASSIFIED	264758
527	80504729 (1053, 1054)	Novel Protein sim. GBank gij73108 [sp]P24215 [XUUA_ECOLI - MANNONATE DEHYDRATASE (D-MANNONATE HYDROLASE)]			UNCLASSIFIED	284769
528	85484134 (1055, 1056)	Novel Protein sim. GBank gij73108 [sp]P24215 [XUUA_ECOLI - MANNONATE DEHYDRATASE (D-MANNONATE HYDROLASE)]			UNCLASSIFIED	58182575, 285017, 285018
529	17936810 (1057, 1058)	Novel Protein sim. GBank gij42144 [emb]CAA232001 - (X00513) NusA protein (nusA) [Escherichia coli]			hydrolase	285019
530	10887336 (1059, 1060)	Novel Protein sim. GBank gij42144 [emb]CAA232001 - (X00513) NusA protein (nusA) [Escherichia coli]			UNCLASSIFIED	264687
531	80228576 (1061, 1062)	Novel Protein sim. GBank gij5282640 [emb]CAB45758.1 - (AL080170) hypothetical protein [Homo sapiens]	Contains protein domain (PF00622) - SPRY domain		UNCLASSIFIED	264555, 264556, 264557, 264558, 18108385
532	80933444 (1063, 1064)	Novel Protein sim. GBank gij5282640 [emb]CAB45758.1 - (AL080170) hypothetical protein [Homo sapiens]			UNCLASSIFIED	264488, 264489, 264258, 264592, 264760, 265021, 264890, 283976, 264558

533	67781531 (1065, 1066)	Novel Protein sim. GBank gi 488363 gb AAD31583.1 AF112299 integral inner nuclear membrane protein MAN1 [Homo sapiens]			284807, 284809, 284786, 35695917, 284630, 284555
534	62388284 (1067, 1068)	Novel Protein sim. GBank gi 2895352 emb CAA04606.1 - (A.001208) pep1 [Streptomyces coelicolor]	UNCLASSIFIED		284905, 285011, 284601, 284602, 284605, 284762, 284786, 285020, 284693, 284838
535	70841850 (1069, 1070)	Novel Protein sim. GBank gi 3878636 emb CAA68953 - (Z49128) similar to cAMP-dependant protein kinase: cDNA EST EMBL: T00719 comes from this gene; cDNA EST yk465d6.3 comes from this gene; cDNA EST yk465d6.5 comes from this gene; cDNA EST yk492f4.3 comes from this gene; cDNA EST y...	Contains protein domain (PF00069) - Eukaryotic protein kinase domain		284906
538	79907207 (1071, 1072)	Novel Protein sim. GBank gi 2495828 sp P55757 YOH1_ SERMA - HYPOTHETICAL 10.1 KD PROTEIN IN BIOA 5'REGION	reductase		18108376, 264805, 264806, 264807, 264809
537	94147446 (1073, 1074)	Novel Protein sim. GBank			265006, 284605, 55274791
536	67821963 (1075, 1076)	Novel Protein sim. GBank gi 134920 sp P21897 SSGP_VOLCA - SULFATED SURFACE GLYCOPROTEIN 185 (SSG 185)	collagen		28331622, 28331624, 28331625, 29331826, 29331827, 264808, 52644045, 33657402, 265017, 264782, 264883, 264288, 264885, 21906765, 35695763, 264558, 60170394, 284559, 22279002
539	28398269 (1077, 1078)	Novel Protein sim. GBank gi 2488433 sp Q12341 HAT1_ YEAST - HISTONE ACETYLTRANSFERASE	histone		284602, 265019
540	79637077 (1079, 1080)	Novel Protein sim. GBank			264893
541	67762266 (1081, 1082)	Novel Protein sim. GBank gi 3862241 db BA34460.1 - (AB018303) KUA0760 protein [Homo sapiens]	transcription factor		18108594, 22276997, 22278998, 264259, 264112, 265009, 33657402, 55812038, 52846317, 265017, 21906785, 264893, 55811578, 264835, 56526486, 264568
542	95295836 (1083, 1084)	Novel Protein sim. GBank gi 5042272 emb CA844526.1 - (AL078616) nuoF, NADH dehydrogenase subunit [Streptomyces coelicolor]	dehydrogenase		264910, 265016, 264889, 264636, 264486
543	79795290 (1085, 1086)	Novel Protein sim. GBank	UNCLASSIFIED		264602, 264906
544	20437191 (1087, 1088)	Novel Protein sim. GBank gi 2791398 emb CAA15994 - (AL021184) hypothetical protein Rv1464 [Mycobacterium tuberculosis]	UNCLASSIFIED		264605
545	80434504 (1089, 1090)	Novel Protein sim. GBank			264786, 264634, 284907, 284592, 264909
546	80249016 (1091, 1092)	gi 468721 gb AAD32237.1 AF14744 - (AF147449) penicillin binding protein 1B [Pseudomonas aeruginosa]			264600, 264602, 21906765
547	11077563 (1093, 1094)	Novel Protein sim. GBank gi 1350855 sp P19176 RPOC_PSEPU - DNA-DIRECTED RNA POLYMERASE BETA' CHAIN (TRANSCRIPTASE BETA' CHAIN) (RNA POLYMERASE BETA' SUBUNIT)	mapolymerase		264604
546	82114836 (1095, 1096)	Novel Protein sim. GBank gi 2330021 AF019250 - kinesin- related protein; KRP; Costal2 [Drosophila melanogaster]	UNCLASSIFIED		264468, 264905, 264910, 264760, 264693, 264639, 264563, 264564

549	95421904 (1097, 1098)	Novel Protein sim. GBank gi 4337460 gb AAD181331 - (AF058195) neuroblastoma-amplified protein [Homo sapiens]		UNCLASSIFIED	264488, 65274572, 18106398, 22278995, 22278998, 22278997, 22278998, 22278999, 264259, 29331824, 66714117, 29331825, 29331826, 35896052, 265007, 265008, 264910, 264592, 33657402, 33109954, 265017, 285018, 285019, 18108351, 264448, 264784, 284369, 264288, 264786, 264688, 264688, 21908765, 21908768, 21908767, 21908768, 21908769, 265020, 264691, 33857023, 264692, 264693, 65274820, 52645129, 33657109, 27488261, 27488262, 27488284, 33657349, 55611576, 18108387, 60432113, 22279002, 264688
550	10886816 (1099, 1100)			UNCLASSIFIED	264688
551	80439980 (1101, 1102)	Novel Protein sim. GBank gi 3122893 sp P94985 SYFB_MYCTU - PHENYLALANYL-TRNA SYNTHETASE BETA CHAIN (PHENYLALANINE--TRNA LIGASE BETA CHAIN) (PHERS)		UNCLASSIFIED	264808, 284809, 264768
552	94672870 (1103, 1104)			UNCLASSIFIED	264889, 264639, 264563
553	80108002 (1105, 1106)	Novel Protein sim. GBank gi 552087 (M33753) - crumbs protein [Drosophila melanogaster]	Contains protein domain (PF00006) - EGF-like domain	glycoprotein	55811957, 264626
554	78618378 (1107, 1108)	Novel Protein sim. GBank gi 501977 gb AAD37857.1 AF13326 - (AF133263) histidine protein kinase-response regulator hybrid protein CvgSY [Pseudomonas syringae pv. syringae]		kinase	264808
555	78968347 (1109, 1110)	Novel Protein sim. GBank gi 131515 sp P02908 PTGA_SALTY - PTS SYSTEM, GLUCOSE-SPECIFIC IIA COMPONENT (EIIA-GLC) (GLUCOSE-PERMEASE IIA COMPONENT) (PHOSPHOTRANSFERASE ENZYME II, A COMPONENT) (EIIA-GLC)	Contains protein domain (PF00356) - phosphoenolpyruvate-dependent sugar phosphotransferase system, EIIA 1	transport	264762
556	20457127 (1111, 1112)	Novel Protein sim. GBank gi 3914014 sp P96380 MFD_MYCTU - TRANSCRIPTION-REPAIR COUPLING FACTOR (TRCF)		transcriptifactor	264508, 264805, 264559
557	19523405 (1113, 1114)	Novel Protein sim. GBank gi 5042273 emb CAB44527.1 - (AL078618) nuOE, NADH dehydrogenase subunit [Streptomyces coelicolor]		dehydrogenase	264488
558	20724429 (1115, 1116)	Novel Protein sim. GBank gi 1170833 sp P45331 METE_HAEN - 5-METHYL-TETRAHYDROPTEROYL-TRIGLUTAMATE-HOMOGYSTEINE METHYLTRANSFERASE (METHIONINE SYNTHASE, VITAMIN-B12 INDEPENDENT ISOZYME) (COBALAMIN-INDEPENDENT METHIONINE SYNTHASE)		UNCLASSIFIED	264602
559	60084353 (1117, 1118)	Novel Protein sim. GBank gi 4980567 gb AAD35173.1 AE00169 - (AE001694) iron(III) ABC transporter, permease protein [Thermotoga maritima]		UNCLASSIFIED	264634

560	80066533 (1118, 1120)	Novel Protein sim. GBank glj2492595[sp]Q53183Y4TR_RHISN - PROBABLE PEPTIDE ABC TRANSPORTER ATP-BINDING PROTEIN Y4TR	Contains protein domain (PF000005) - ABC transporter	transport	18108398, 264908, 264602, 264604, 18108374
561	20293187 (1121, 1122)			UNCLASSIFIED	264600
562	11898181 (1123, 1124)			UNCLASSIFIED	264689
563	79781420 (1125, 1126)	Novel Protein sim. GBank glj4104925 (AF042276) - poly(hydroxycarboxylate) granule associated protein GA2 [Pseudomonas putida]		UNCLASSIFIED	264910, 264691
564	56716390 (1127, 1128)	Novel Protein sim. GBank glj2792310 (AF040570) - unknown [Amycolatopsis mediterranei]		dehydrogenase	264592
565	56485618 (1129, 1130)	Novel Protein sim. GBank glj3448294[dbj]BAA324621 - (AB011532) MEGF8 [Rattus norvegicus]	Contains protein domain (PF00008) - EGF-like domain	synthase	265010
566	94323868 (1131, 1132)	Novel Protein sim. GBank glj4539598[emb]CAB38467.11 - (AL035636) putative helicase [Streptomyces coelicolor]		helicase	264909, 264510, 265008, 264910, 264758, 264800, 264602, 264604, 264605, 264788, 264687, 264689, 35895917, 264693, 65274620, 264488
567	79560955 (1133, 1134)	Novel Protein sim. GBank glj100506[sp]S17455 - Malate dehydrogenase (oxaloacetate-decarboxylating) (NADP+) (EC 1.1.1.40) - Flaveria trinervia (fragment)	Contains protein domain (PF00390) - Malic enzyme	UNCLASSIFIED	264681, 264691, 264593
568	94681793 (1135, 1136)	Novel Protein sim. GBank glj3915843[sp]O31212[RS2_STRCO - 30S RIBOSOMAL PROTEIN S2	Contains protein domain (PF00318) - Ribosomal protein S2	UNCLASSIFIED	264585
569	39506897 (1137, 1138)			transport	18108378, 18108387, 264585
570	78375927 (1139, 1140)	Novel Protein sim. GBank glj11512[sp]P21627[BRAD_PSEAE - HIGH-AFFINITY BRANCHED-CHAIN AMINO ACID TRANSPORT PROTEIN BRAD		UNCLASSIFIED	264907, 264809
571	79783881 (1141, 1142)			UNCLASSIFIED	264782
572	38986838 (1143, 1144)	Novel Protein sim. GBank glj4539223[emb]CAB39881.11 - (AL049497) putative integral membrane protein [Streptomyces coelicolor]		UNCLASSIFIED	265007, 264601
573	20715521 (1145, 1146)			UNCLASSIFIED	264638
574	13521592 (1147, 1148)	Novel Protein sim. GBank glj118784[sp]P10443[DP3A_ECOLI - DNA POLYMERASE III, ALPHA CHAIN		polymerase	264687
575	13078418 (1149, 1150)				264605
576	20482248 (1151, 1152)	Novel Protein sim. GBank glj545782[emb]CAB49116.11 - (AJ248283) PAB2227 [Pyrococcus abyssi]	Contains protein domain (PF00346) - Respiratory-chain NADH dehydrogenase, 49 Kd subunit	dehydrogenase	35898052, 264636
577	66727102 (1153, 1154)	Novel Protein sim. GBank glj5042274[emb]CAB44528.11 - (AL078616) nuoD, NADH dehydrogenase subunit [Streptomyces coelicolor]			264638
578	11804477 (1155, 1156)	Novel Protein sim. GBank glj1723061[sp]Q11046[Y089_MYCTU - HYPOTHETICAL ABC TRANSPORTER ATP-BINDING PROTEIN CY50.08		transport	264682, 264556
579	11784723 (1157, 1158)				

580	80059417 (1159, 1160)				22278899, 35696052, 284555, 284558, 284556
581	79230033 (1181, 1182)	Novel Protein sim. GBank gl 3243131 (AF045777) - Itin	UNCLASSIFIED		285008, 284564
582	80049617 (1183, 1184)	[Drosophila melanogaster]	struct	Contains protein domain (PF00047) - Immunoglobulin domain	285021, 284555, 284557
583	79321392 (1185, 1186)	Novel Protein sim. GBank gl 2501182sp P77728 YAJR_ECOLI - HYPOTHETICAL 49.0 KD PROTEIN IN ABPA-CYOE INTERGENIC REGION	transport		284594
584	79845024 (1187, 1188)	Novel Protein sim. GBank gl 3082221 db BAA34470.1 - (AB018283) KIAA0750 protein [Homo sapiens]	UNCLASSIFIED		284488, 284906, 284788, 284887, 35696423
585	79561454 (1189, 1190)	Novel Protein sim. GBank gl 3082221 db BAA34470.1 - (AB018283) KIAA0750 protein [Homo sapiens]	UNCLASSIFIED		285018, 284684, 21908789
586	38277486 (1171, 1172)	Novel Protein sim. GBank gl 4467250 emb CAB37575 - (AL035569) probable Glu-RNA Gln amidotransferase subunit [Streptomyces coelicolor]	UNCLASSIFIED		284908, 285007
587	80497359 (1173, 1174)	Novel Protein sim. GBank gl 5689519 db BAA83043.1 - (AB029014) KIAA1091 protein [Homo sapiens]	hydrolase		284600, 284602, 284605, 284768, 284690, 284557
588	79857239 (1175, 1176)	Novel Protein sim. GBank gl 5689519 db BAA83043.1 - (AB029014) KIAA1091 protein [Homo sapiens]	UNCLASSIFIED		285020, 284892
589	79805828 (1177, 1178)		UNCLASSIFIED		22278899, 284907, 284909, 284510, 285009, 285010, 284887, 284789, 35695917, 18108376, 284634, 284636, 284838
590	79815829 (1179, 1180)	Novel Protein sim. GBank gl 2143293 emb CAB09390 - (Z95972) pncB [Mycobacterium tuberculosis]	UNCLASSIFIED		284908, 284909
591	10313540 (1181, 1182)	Novel Protein sim. GBank gl 2143293 emb CAB09390 - (Z95972) pncB [Mycobacterium tuberculosis]	mapolymerase		284891
592	13889767 (1183, 1184)	Novel Protein sim. GBank gl 4511983 gb AAD21543.1 - (AF088896) electrotransfer ubiquinone oxidoreductase [Zymomonas mobilis]	MHC		263972
593	82348888 (1185, 1186)	Novel Protein sim. GBank gl 4511983 gb AAD21543.1 - (AF088896) electrotransfer ubiquinone oxidoreductase [Zymomonas mobilis]	dehydrogenase		284511, 284782, 284789, 284488
594	20212392 (1187, 1188)	Novel Protein sim. GBank gl 1272388 (U51699) - LigE [Vibrio parahaemolyticus]	UNCLASSIFIED		284605
595	10064064 (1189, 1190)	Novel Protein sim. GBank gl 131490 sp P20966 PTFB_ECOLI - PTS SYSTEM, FRUCTOSE-SPECIFIC IIBC COMPONENT (EIIBC-FRU) (FRUCTOSE-PERMEASE IIBC COMPONENT) (PHOSPHOTRANSFERASE ENZYME II, BC COMPONENT) (EIIF-FRU)			284768
596	13085170 (1191, 1192)		UNCLASSIFIED		284636
597	80259003 (1193, 1194)		UNCLASSIFIED		284592
598	94140216 (1195, 1196)		UNCLASSIFIED		284758, 55810764, 284555, 284556, 284637, 83373044
599	20385137 (1197, 1198)	Novel Protein sim. GBank gl 125329 sp P04951 KDSB_ECOLI - 3-DEOXY-MANNO-OCTULONATE CYTIDYLTRANSFERASE (CMP-KDO SYNTHETASE) (CMP-2-KETO-3-DEOXYOCTULONIC ACID SYNTHETASE) (CKS)	UNCLASSIFIED		284603
600	10357663 (1199, 1200)				
601	79610404 (1201, 1202)	Novel Protein sim. GBank gl 2127414 pr S60084 - hypothetical protein 2 - Corynebacterium glutamicum	UNCLASSIFIED		284808, 284510

602	79250602 (1203, 1204)	Novel Protein sim. GBank g 352288 gb AAC34243.1 - (AC004411) putative pto kinase [Arabisopsis thaliana]	Contains protein domain (PF00069) - Eukaryotic protein kinase domain	kinase	265007
603	11466067 (1205, 1206)			UNCLASSIFIED	264595
604	81875420 (1207, 1208)			UNCLASSIFIED	264758
605	20436657 (1208, 1210)	Novel Protein sim. GBank g 11753222 sp P44917 Y883_HAEN - HYPOTHETICAL PROTEIN H0883		UNCLASSIFIED	264605
608	80334582 (1211, 1212)	Novel Protein sim. GBank g 5020284 gb AAD38043.1 AF15135 - (AF151383) Cdc42 GTPase-activating protein [Mus musculus]		UNCLASSIFIED	264764
607	95381508 (1213, 1214)	Novel Protein sim. GBank g 188864 N74027 - mucin [Homo sapiens]		UNCLASSIFIED	264508, 264908, 65658542, 264682, 264687, 264689, 264534, 18108378, 35898423, 264638, 264555, 264638
608	11810888 (1215, 1216)			UNCLASSIFIED	264682
609	80064775 (1217, 1218)	Novel Protein sim. GBank g 248870 sp P55552 Y4LL_RHISN - HYPOTHETICAL 91.8 KD PROTEIN Y4LL	Contains protein domain (PF00988) - PAS domain	UNCLASSIFIED	264605
610	79828413 (1218, 1220)				264692
611	87588205 (1221, 1222)				264508, 264905, 264907, 264908, 264909, 264511, 264910, 264758, 264604, 264684, 264768, 264889, 264692, 264628, 264635, 264636, 264837, 264558
612	95287851 (1223, 1224)	Novel Protein sim. GBank g 1877366 emb CAB07118 - (Z92772) recD [Mycobacterium tuberculosis]	Contains protein domain (PF01443) - Viral (Superfamily 1) RNA helicase	nuclease	264600, 264801, 264604, 264769, 264558, 264565
613	7523475 (1225, 1226)			UNCLASSIFIED	264389
614	79989348 (1227, 1228)	Novel Protein sim. GBank g 5114231 gb AAD40238.1 AF13870 - (AF138709) histidine kinase YycG [Staphylococcus aureus]		kinase	18108372, 264583
615	39586996 (1229, 1230)	Novel Protein sim. GBank g 1339950 gb BA12741 - (D85230) large subunit of NADH-dependent glutamate synthase [Plectonema boryanum]		synthase	264600, 264602, 264628
616	20485331 (1231, 1232)	Novel Protein sim. GBank g 544387 sp P35673 GALE_ERWAM - UDP-GLUCOSE 4-EPIMERASE (GALACTOWALDENASE) (UDP-GALACTOSE 4-EPIMERASE)		isomerase	264605
617	91227222 (1233, 1234)	Novel Protein sim. GBank g 2488087 sp Q60789 TNF3_MOUSE - TUMOR NECROSIS FACTOR, ALPHA-INDUCED PROTEIN 3 (PUTATIVE DNA BINDING PROTEIN A20) (ZINC FINGER PROTEIN A20)	Contains protein domain (PF00641) - Zn-finger in Ran binding protein and others.	Inf	52845158, 21808785, 35698423, 21806768, 21808769, 22278994, 35698288, 22278996, 285020, 265021, 265007, 265008, 264638, 52844150, 33657023, 264692, 264693, 29331822, 29331824, 55812038, 83373044, 56182181, 80424268, 68714117, 29331825, 33657109, 29331828, 33657182, 29331827, 35698052, 29331828, 27486282, 33657349, 55528488, 285018, 265019, 22279002, 264482, 284448, 29331830, 68712502, 264909

618	20632843 (1235, 1236)	Novel Protein sim. GBank gl 5459388 emb CAB50746.1 - (AL096839) putative aminotransferase [Streptomyces coelicolor]		isomerase	264603
619	91227224 (1237, 1238)				56984075, 28331828, 33656970, 265008, 33657402, 33109954, 87168559, 264448, 18108374, 83373044
820	81163143 (1239, 1240)	Novel Protein sim. GBank gl 484335 sp Q05822 DUS2_MOUSE - DUAL SPECIFICITY PROTEIN PHOSPHATASE 2 (DUAL SPECIFICITY PROTEIN PHOSPHATASE PAC-1)		phosphatase	28146498, 264758, 264369, 28148627
621	80239251 (1241, 1242)	Novel Protein sim. GBank gl 2633557 emb CAB13060 - (Z89110) yJdF [Bacillus subtilis]		UNCLASSIFIED	264556, 284558, 264639
622	20456427 (1243, 1244)	Novel Protein sim. GBank gl 1857710 gb AAB48482 - (U87224) contactin associated protein [Rattus norvegicus]		UNCLASSIFIED	284605
623	10131798 (1245, 1246)	Novel Protein sim. GBank gl 1705703 sp P52225 CCMF_PSEFL - CYTOCHROME C-TYPE BIOGENESIS PROTEIN CYCK	Contains protein domain (PF000054) - Laminin G domain	laminin	264906
624	18534127 (1247, 1248)	Novel Protein sim. GBank gl 1705703 sp P52225 CCMF_PSEFL - CYTOCHROME C-TYPE BIOGENESIS PROTEIN CYCK		cytochrome	264596
825	13084619 (1249, 1250)	Novel Protein sim. GBank gl 2894252 emb CAA17114.1 - (AL021841) hypothetical protein Rv3342 [Mycobacterium tuberculosis]		UNCLASSIFIED	264688
626	68062603 (1251, 1252)	Novel Protein sim. GBank gl 16592 sp P33233 AGA1_YEAST - A-AGGLUTININ ATTACHMENT SUBUNIT PRECURSOR		UNCLASSIFIED	28331822, 264905, 264808, 33657023, 33657109, 264558
627	80255457 (1253, 1254)	Novel Protein sim. GBank gl 3098419 (AF040944) - P140 [Mus musculus]		UNCLASSIFIED	18108394, 264112, 264593, 265022, 264635
628	80077096 (1255, 1256)	Novel Protein sim. GBank gl 1711543 sp P40526 SSP1_SCHPO - SERINE/THREONINE-PROTEIN KINASE SSP1	Contains protein domain (PF00069) - Eukaryotic protein kinase domain	kinase	264600
829	79851802 (1257, 1258)	Novel Protein sim. GBank gl 11433204 (U34305) - ORF2. Method: conceptual translation supplied by author. [Shigella sonnei]		isomerase	264906, 264907
630	38565156 (1259, 1260)	Novel Protein sim. GBank gl 3236368 (AF064748) - S3-12 [Mus musculus]		UNCLASSIFIED	264480
631	20598716 (1261, 1262)	Novel Protein sim. GBank gl 140867 sp P11688 YGG8_ECOLI - HYPOTHETICAL 30.9 KD PROTEIN IN SBM-FBA INTERGENIC REGION (ORF 4) (F286)			263978
632	27843890 (1263, 1264)			UNCLASSIFIED	264906, 264600, 264605, 264769, 264889, 264486
633	80477772 (1265, 1266)			UNCLASSIFIED	264769
634	17938608 (1267, 1268)			UNCLASSIFIED	265019
635	79574506 (1269, 1270)			UNCLASSIFIED	264689
636	79910981 (1271, 1272)			UNCLASSIFIED	284598, 264762, 264693

837	62455706 (1273, 1274)	Novel Protein sim. GBank gij2326739[embjCAB10953] - (Z98266) recN [Mycobacterium tuberculosis]		nuclease	264906, 264907, 264510, 264511, 264601, 264602, 264603, 264604, 264605, 18108351, 264782, 264786, 264687, 264789, 264688, 35695917, 264883, 264834, 264838, 264639, 264559, 18108385
838	14987457 (1275, 1276)	Novel Protein sim. GBank gij4678662[embjCAB41074.1] - (AL049645) putative large ATP-binding protein [Streptomyces coelicolor]			264636
839	80204210 (1277, 1278)	Novel Protein sim. GBank gij459828[dbjBAA76836.1] - (AB023209) KIAA0992 protein [Homo sapiens]		si/ruct	264112, 263874
840	17929579 (1279, 1280)	Novel Protein sim. GBank gij1432063 (U60981) - homolog to Skp1p, an evolutionarily conserved kinetochore protein in budding yeast [Arabidopsis thaliana]	Contains protein domain (PF01466) - Skp1 family	mapolymerase	265009, 265010
841	79636398 (1281, 1282)			UNCLASSIFIED	264693
842	18896737 (1283, 1284)			UNCLASSIFIED	264565
843	81516220 (1285, 1286)			UNCLASSIFIED	264906, 264908, 264758, 264288, 264632, 264635, 264639, 264584
844	11751367 (1287, 1288)			UNCLASSIFIED	264884
845	95010907 (1289, 1290)			UNCLASSIFIED	264906, 264762, 264693, 264639, 264559
846	80069083 (1291, 1292)			UNCLASSIFIED	264595, 264566
847	80257085 (1293, 1294)	Novel Protein sim. GBank gij4507813[refNP_003738.1]pTNKS - TANKYRASE	Contains protein domain (PF00023) - Ank repeat	transcriptfador	264909, 264591
848	80077428 (1295, 1296)	Novel Protein sim. GBank gij1044063[bbbs]169646 - prolamine [Monodonla turbinata, gonads, Peptide, 106 aa]		UNCLASSIFIED	264600
849	80247447 (1297, 1298)			UNCLASSIFIED	263978
850	11798318 (1299, 1300)			UNCLASSIFIED	264886
851	11776932 (1301, 1302)	Novel Protein sim. GBank gij1346916[spjP12283]PURA_ECOLI - ADENYLOSUCCINATE SYNTHETASE (IMP-ASPARTATE LIGASE)		UNCLASSIFIED	264602, 264638
852	85516704 (1303, 1304)			UNCLASSIFIED	264905, 264907, 264909, 263978, 264637
853	82124947 (1305, 1306)	Novel Protein sim. GBank gij172297[spjQ10538]Y03C_MYCTU - HYPOTHETICAL 82.8 KD PROTEIN CY130.12C		UNCLASSIFIED	22276996, 264510, 264511, 264512, 264593, 21906754, 264603, 264760, 18106376, 264556
854	95010569 (1307, 1308)			UNCLASSIFIED	264906, 264595, 264632
855	79320692 (1309, 1310)	Novel Protein sim. GBank gij130327[spjP26647]PLSC_ECOLI - 1-ACYL-SN-GLYCEROL-3-PHOSPHATE ACYLTRANSFERASE (1-AGP ACYLTRANSFERASE) (1-AGPAT) (LYSOPHOSPHATIDIC ACID ACYLTRANSFERASE) (LPAAT)	Contains protein domain (PF01553) - Acyltransferase	transferase	264592
856	80416739 (1311, 1312)			UNCLASSIFIED	264602, 264605, 264766, 264691
857	20611010 (1313, 1314)			UNCLASSIFIED	264557, 264558

658	87761815 (1315, 1316)	Novel Protein sim. GBank gl 5689493 dbj BAA83030.1 - (AB029001) KIAA1078 protein [Homo sapiens]		UNCLASSIFIED	22278996, 60432049, 29331822, 29331824, 29331828, 265007, 265009, 33657402, 33857084, 285017, 284448, 21906785, 21906766, 263987, 20281149, 18108370, 18108374, 264482
659	87718663 (1317, 1318)	Novel Protein sim. GBank gl 2137872 prt J48724 - zinc finger protein PZF - mouse	Contains protein domain (PF00086) - Zinc finger, C2H2 type	transcriptfactor	22278998, 60432049, 66714117, 29331827, 285007, 284788, 58181582, 18108359, 18108365, 18108370, 18108381
660	81887922 (1319, 1320)	Novel Protein sim. GBank		UNCLASSIFIED	264757
661	80028023 (1321, 1322)	Novel Protein sim. GBank gl 134180 p P1540 I SACY_BACSU - LEVANSUCRASE AND SUCRASE SYNTHESIS OPERON ANTERMINATOR	Contains protein domain (PF00874) - Transcriptional antiterminator bglG family	UNCLASSIFIED	264510, 265009, 264600, 264602, 284803, 284804, 264605, 32833986, 18108378, 264636, 18108387, 22278000
662	20463731 (1323, 1324)	Novel Protein sim. GBank gl 454528 gb AAD22450.1 AF11618 - (AF116183) SecA homolog [Acetobacterium acetiomycescomitans]		UNCLASSIFIED	264805
663	20628080 (1325, 1326)	Novel Protein sim. GBank gl 5689250 dbj BAA82881.1 - (AB024335) similar to orf5 [Comamonas testosteroni]		dehydrogenase	264605
664	80508512 (1327, 1328)	Novel Protein sim. GBank gl 1652848 dbj BAA17768 - (D90909) DNA photolyase [Synecococcus sp.]		UNCLASSIFIED	264768
665	80078053 (1329, 1330)	Novel Protein sim. GBank gl 116841 p P21640 COBJ_PSEDE - PRECORRIN-3B C17 METHYLTRANSFERASE (PRECORRIN-3 METHYLTRANSFERASE)		isomerase	264600
666	78603142 (1331, 1332)	Novel Protein sim. GBank gl 3261829 emb CAB10927 - (Z98280) hypothetical protein Rv1230c [Mycobacterium tuberculosis]		glycoprotein	264907, 285007
667	94631802 (1333, 1334)	Novel Protein sim. GBank gl 568885 dbj BAA82702.1 - (AB017438) Orf5 [Streptomyces coelicolor]		UNCLASSIFIED	264688, 264602, 264593
668	82051891 (1335, 1336)	Novel Protein sim. GBank gl 3581853 emb CAA20809 - (AL031541) 50S ribosomal protein L20 [Streptomyces coelicolor]	Contains protein domain (PF00453) - Ribosomal protein L20	ribosomalprot	264905, 264906, 264908, 284908, 284600, 264601, 264603, 264605, 264760, 264889, 264838, 264638, 264839
669	12867154 (1337, 1338)	Novel Protein sim. GBank gl 2582531 (AF026444) - 2- isopropylmalate synthase [Streptomyces coelicolor]		UNCLASSIFIED	264837
670	80238549 (1339, 1340)	Novel Protein sim. GBank gl 2582531 (AF026444) - 2- isopropylmalate synthase [Streptomyces coelicolor]		synthase	264905, 264806, 264808, 264601, 264762, 264768, 264889, 284638, 18108385, 264488
671	79601368 (1341, 1342)	Novel Protein sim. GBank gl 2114430 (U92703) - Oil-1/EBF-like-3 transcription factor [Mus musculus]	Contains protein domain (PF00023) - Ank repeat	UNCLASSIFIED	264890, 284692, 264693, 264636, 18108387
672	79834371 (1343, 1344)	Novel Protein sim. GBank gl 2114430 (U92703) - Oil-1/EBF-like-3 transcription factor [Mus musculus]		transcriptfactor	264910, 265017
673	82285798 (1345, 1346)	Novel Protein sim. GBank gl 4589285 gb AAD28430.1 AF13515 - (AF135154) ferric alkaliphilic siderophore receptor [Bordetella pertussis]			284759
674	79198259 (1347, 1348)			UNCLASSIFIED	264629

875	87898870 (1348, 1350)	Novel Protein sim. GBank gi 4980755 gb AAD35347.1 AE00170 - (AE001708) D- alanine-D-alanine ligase [Thermotoga maritima]	Contains protein domain (PF01820) - D-ala D-ala ligase	UNCLASSIFIED	264488, 22278998, 66714117, 264508, 284511, 265008, 60433438, 264600, 284801, 264602, 264603, 264804, 264605, 264762, 264687, 264769, 60431602, 18108374, 264638, 264638 285010
876	78898607 (1351, 1352)	Novel Protein sim. GBank gi 1723566 sp Q10479 YDF7_SCHPO - PUTATIVE GLUCOSYLTRANSFERASE C17C9.07			
877	21644312 (1353, 1354)	Novel Protein sim. GBank gi 887208 (U03976) - dynein heavy chain isotype 5C [Tripeustes gratilis]		ATPase-associated	264591, 264632
878	84225200 (1355, 1356)	Novel Protein sim. GBank gi 1586274 pr 2203365A - laminin alpha5 [Mus musculus]	Contains protein domain (PF00053) - Laminin EGF-like (Domains III and V)	laminin	264758, 264682, 264557
879	79868855 (1357, 1358)	Novel Protein sim. GBank gi 3928723 emb CAA22219 - (A034355) putative ABC transporter [Streptomyces coelicolor]		UNCLASSIFIED	22278998, 264693
880	20726424 (1359, 1360)				
881	94322017 (1361, 1362)	Novel Protein sim. GBank gi 5174483 ref NP_006050.1 pLAC - laminin, gamma 3	Contains protein domain (PF00053) - Laminin EGF-like (Domains III and V)	laminin	264600, 264602 264102, 264907, 264908, 285008, 264893, 263972, 83373044, 264568
882	11392478 (1363, 1364)	Novel Protein sim. GBank	Contains protein domain (PF00762) - Dual specificity phosphatase, catalytic domain	UNCLASSIFIED	264585 264634
883	80083680 (1365, 1366)	gi 4758208 ref NP_004081.1 pDUSP - dual specificity phosphatase 3 (vaccinia virus phosphatase VH1-related)			264605
884	20465367 (1367, 1368)	Novel Protein sim. GBank gi 5420387 emb CA846679.1 - (A243459) proteophosphoglycan [Leishmania major]			
885	80246735 (1369, 1370)				264809, 263967, 263981
886	79208808 (1371, 1372)				264631
887	80085829 (1373, 1374)				264693, 264635
888	79853412 (1375, 1376)	Novel Protein sim. GBank gi 2688962 (AF027768) - LspA [Serratia marcescens]		peptidase	264907, 264638
889	88064258 (1377, 1378)	Novel Protein sim. GBank gi 3046831 (AF048330) - PPAR gamma coactivator [Mus musculus]	Contains protein domain (PF00076) - RNA recognition motif, (a.k.a. RRM, RBD, or RNP domain)	UNCLASSIFIED	264806, 264907, 265007, 265008, 60433438, 21908754, 264760, 18106358, 21808788, 21906769, 265021, 18108381, 263974, 18108378, 264557, 18106385, 22279002 264510, 264511, 264764, 264789
890	80388750 (1379, 1380)	Novel Protein sim. GBank gi 2498941 sp Q15428 SP62_HUMAN - SPLICEOSOME ASSOCIATED PROTEIN 62 (SAP 62) (SF3A88)		UNCLASSIFIED	
891	81654392 (1381, 1382)			UNCLASSIFIED	264757
892	83608926 (1383, 1384)	Novel Protein sim. GBank gi 5420387 emb CA846679.1 - (A243459) proteophosphoglycan [Leishmania major]	Contains protein domain (PF00097) - Zinc finger, C3HC4 type (RING finger)	UNCLASSIFIED	55812038, 55811957, 265018, 55811150, 18108351, 264908, 60431528, 264594
893	79588116 (1385, 1386)	Novel Protein sim. GBank gi 854065 emb CAA58337 - (X63413) U88 [Human herpesvirus 6]		UNCLASSIFIED	264835
894	82455983 (1387, 1388)	Novel Protein sim. GBank gi 267327 sp Q01033 VG48_HSVSA - HYPOTHETICAL GENE 48 PROTEIN			22278998, 264510, 264602, 264603, 264762, 264687, 264769, 264893

695	94147849 (1389, 1390)	Novel Protein sim. GBank glij4468339[embj CA838059.1] - (AJ010901) MUC4 [Homo sapiens]	Contains protein domain (PF00094) - von Willebrand factor type D domain	UNCLASSIFIED	56182575, 264509, 264905, 264907, 29331830, 264908, 264511, 285007, 264910, 284758, 264764, 264288, 85274791
696	79830882 (1391, 1392)	Novel Protein sim. GBank glij2649950 (AE001058) - glutamine ABC transporter, ATP-binding protein (glnC) [Archaeoglobus fulgidus]	Contains protein domain (PF00005) - ABC transporter	transport	264905, 264595
697	11787889 (1393, 1394)	Novel Protein sim. GBank glij1731343[spj Q10694]YY25_MYCTU - HYPOTHETICAL 24.4 KD PROTEIN CY49.25	Contains protein domain (PF01838) - Transposase	UNCLASSIFIED	264882
698	66695862 (1395, 1396)			UNCLASSIFIED	264888, 35895917
699	79582558 (1397, 1398)			UNCLASSIFIED	264682
700	79639098 (1399, 1400)			UNCLASSIFIED	264693
701	80230242 (1401, 1402)	Novel Protein sim. GBank glij1001236[dbj BAA10471] - (D64003) hypothetical protein [Synachocystis sp.]		UNCLASSIFIED	264488, 264510, 264511, 264602, 264605, 284688
702	79814789 (1403, 1404)	Novel Protein sim. GBank glij2488835[spj Q46338]SOXG_CORSP - SARCOSINE OXIDASE GAMMA SUBUNIT		UNCLASSIFIED	264809
703	20448820 (1405, 1406)	Novel Protein sim. GBank glij3150513 (AF087219) - contains similarity to the kelch/MIPP family [Caenorhabditis elegans]		oxidase	264604
704	94312224 (1407, 1408)	Novel Protein sim. GBank glij3150513 (AF087219) - contains similarity to the kelch/MIPP family [Caenorhabditis elegans]	Contains protein domain (PF01344) - Kelch motif	UNCLASSIFIED	264288, 58181562, 33657109, 264629, 55811578
705	17932141 (1409, 1410)	Novel Protein sim. GBank glij421091[pirj S30730] - hypothetical protein o206 - Escherichia coli		UNCLASSIFIED	265008
706	20288062 (1411, 1412)	Novel Protein sim. GBank glij3024872[spj Q5780]Y074_SYNY3 - HYPOTHETICAL 52.8 KD PROTEIN SLR0074			264600
707	20838085 (1413, 1414)	Novel Protein sim. GBank glij3420608[gbj AAC31907.1] - (AF075709) ABC transporter ATP-binding subunit [Pseudomonas putida]		transport	264603
708	20708292 (1415, 1416)	Novel Protein sim. GBank glij3649741[embj CAA03985] - (AJ000281) mucin [Homo sapiens]			264801, 264892
709	88001439 (1417, 1418)	Novel Protein sim. GBank glij3080425[embj CAA18744.1] - (AL022804) putative protein [Arabidopsis thaliana]		strud	18108398, 264637, 264908, 264909
710	11356683 (1419, 1420)	Novel Protein sim. GBank glij4758688[refj NP_002323.1]pLRP1 - low density lipoprotein related protein 1 (alpha-2-macroglobulin receptor)			264389
711	17931418 (1421, 1422)	Novel Protein sim. GBank glij1703288[spj Q11056]AM12_MYCTU - PUTATIVE AMIDASE CY50.19C	Contains protein domain (PF00058) - Low-density lipoprotein receptor repeat class B		265019
712	80258164 (1423, 1424)	Novel Protein sim. GBank glij4502351[refj NP_001692.1]pBAAT - bile acid Coenzyme A: amino acid N-acyltransferase: glycine N-choloyltransferase		apolipoprotein	264581
713	79263128 (1425, 1426)	Novel Protein sim. GBank glij1703288[spj Q11056]AM12_MYCTU - PUTATIVE AMIDASE CY50.19C		hydrolase	264908, 264907
714	27647851 (1427, 1428)	Novel Protein sim. GBank glij4502351[refj NP_001692.1]pBAAT - bile acid Coenzyme A: amino acid N-acyltransferase: glycine N-choloyltransferase			264508, 264555

715	78639423 (1428, 1430)	Novel Protein sim. GBank gl 1789035 (AE000352) - ori, hypothetical protein [Escherichia coli]		UNCLASSIFIED	264907
716	70550072 (1431, 1432)	Novel Protein sim. GBank gl 2494074 sp P55653 GABD_RHISN - PROBABLE SUCCINATE-SEMIALDEHYDE DEHYDROGENASE (NADP+)(SSDH)		dehydrogenase	264632
717	70491842 (1433, 1434)	Novel Protein sim. GBank gl 3873679 emb CAA94886 (Z71178) similar to pro-collagen domains; cDNA EST EMBL:027978 comes from this gene; cDNA EST EMBL:027977 comes from this gene; cDNA EST EMBL:034199 comes from this gene; cDNA EST EMBL:064392 comes from this gene; cDNA EST EMBL:02104302 emb CAB08631 - (Z95387) hypothetical protein Rv2611c [Mycobacterium tuberculosis]	Contains protein domain (PF00093) - von Willebrand factor type C domain	kinase	18108392, 22278894, 22278998, 265008, 265018, 264881, 18108354, 264884, 264885, 264886, 264887, 264889, 21906769, 18108361, 264691, 264692, 55810764, 264835, 18108381, 18108382, 83373044, 18108388
718	17679584 (1437, 1438)	Novel Protein sim. GBank gl 2104302 emb CAB08631 - (Z95387) hypothetical protein Rv2611c [Mycobacterium tuberculosis]		UNCLASSIFIED	265011
719	79841684 (1439, 1440)	Novel Protein sim. GBank gl 123530 sp P04929 HRPX_PLALO - HISTIDINE-RICH GLYCOPROTEIN PRECURSOR			264908
720	15020160 (1441, 1442)	Novel Protein sim. GBank gl 498253 (U02372) - integrase [Vibrio cholerae]		UNCLASSIFIED	264629
721	9862603 (1443, 1444)	Novel Protein sim. GBank gl 2253054 emb CAB10705 - (Z97559) hypothetical protein Rv2114 [Mycobacterium tuberculosis]		UNCLASSIFIED	264810
722	19755599 (1445, 1446)	Novel Protein sim. GBank gl 4063015 (AF083061) - protease PrtA [Pseudomonas fluorescens]	Contains protein domain (PF00353) - Hemolysin-type calcium-binding proteins	protease	264908
723	10126484 (1447, 1448)	Novel Protein sim. GBank gl 2633810 emb CAB13411 - (Z99112) similar to hypothetical proteins [Bacillus subtilis]		UNCLASSIFIED	264905, 264807
724	7878678 (1449, 1450)	Novel Protein sim. GBank gl 2494074 sp P55653 GABD_RHISN - PROBABLE SUCCINATE-SEMIALDEHYDE DEHYDROGENASE (NADP+)(SSDH)		UNCLASSIFIED	264636
725	13086282 (1451, 1452)	Novel Protein sim. GBank gl 1146182 (L47838) - putative [Bacillus subtilis]		UNCLASSIFIED	264634
726	13522872 (1453, 1454)	Novel Protein sim. GBank gl 1146182 (L47838) - putative [Bacillus subtilis]			264567
727	20269471 (1455, 1456)	Novel Protein sim. GBank gl 2494074 sp P55653 GABD_RHISN - PROBABLE SUCCINATE-SEMIALDEHYDE DEHYDROGENASE (NADP+)(SSDH)		UNCLASSIFIED	264490
728	11293753 (1457, 1458)	Novel Protein sim. GBank gl 2494074 sp P55653 GABD_RHISN - PROBABLE SUCCINATE-SEMIALDEHYDE DEHYDROGENASE (NADP+)(SSDH)		isomerase	264564
729	19500373 (1459, 1460)	Novel Protein sim. GBank gl 1146182 (L47838) - putative [Bacillus subtilis]		UNCLASSIFIED	264605
730	80058750 (1461, 1462)	Novel Protein sim. GBank gl 1146182 (L47838) - putative [Bacillus subtilis]		struct	264591, 264594, 264595
731	80258175 (1463, 1464)	Novel Protein sim. GBank gl 1146182 (L47838) - putative [Bacillus subtilis]		UNCLASSIFIED	264804
732	20468839 (1465, 1466)	Novel Protein sim. GBank gl 3184080 emb CAA19336 - (AL023781) hypothetical protein [Schizosaccharomyces pombe]		UNCLASSIFIED	264604
733	20435987 (1467, 1468)	Novel Protein sim. GBank gl 3184080 emb CAA19336 - (AL023781) hypothetical protein [Schizosaccharomyces pombe]		ubiquitin	

735	11607959 (1489, 1470)	Novel Protein sim. GBank gi 401582 sp p27432 YICE_ECOLI - HYPOTHETICAL 48.9 KD PROTEIN IN GLTS-SELC INTERGENIC REGION			264594
736	10879734 (1471, 1472)	Novel Protein sim. GBank gi 400831 sp p31135 POTH_ECOLI - PUTRESCINE TRANSPORT SYSTEM PERMEASE PROTEIN POTH	transport	Contains protein domain (PF00528) - Binding-protein-dependent transport systems Inner membrane component	264836
737	78945340 (1473, 1474)		UNCLASSIFIED	Contains protein domain (PF00815) - Regulator of G protein signaling domain	265020
738	17995353 (1475, 1478)				265008
739	79833870 (1477, 1478)	Novel Protein sim. GBank gi 2506867 sp p33225 TORA_ECOLI - TRIMETHYLAMINE- N-OXIDE REDUCTASE PRECURSOR (TMAO REDUCTASE) (TRIMETHYLAMINE OXIDASE)	oxidase		264810
740	19881557 (1479, 1480)				264907, 264764, 264634, 264637
741	79827273 (1481, 1482)	Novel Protein sim. GBank gi 3261828 emb CAB10925 - (Z98280) mmp [Mycobacterium tuberculosis]	UNCLASSIFIED	Contains protein domain (PF01683) - Domain of unknown function	264689, 35698286, 264510, 264808, 18108362
742	82393795 (1483, 1484)	Novel Protein sim. GBank gi 3877494 emb CAA88472.1 - (Z48593) ATP binding protein with similarity to the CDC48/PAS1/SEC18 family; cDNA EST EMBL:D65037 comes from this gene; cDNA EST EMBL:D68340 comes from this gene; cDNA EST EMBL:D65048 comes from this gene; cDNA EST EMBL:D6845...	UNCLASSIFIED		29331822, 264910, 264762
743	82300051 (1485, 1486)	Novel Protein sim. GBank gi 127420 sp p19888 MTBA_BACAR - MODIFICATION METHYLASE BANI (CYTOSINE-SPECIFIC METHYLTRANSFERASE BANI) (MBANI)	UNCLASSIFIED	Contains protein domain (PF00145) - C-5 cytosine-specific DNA methylase	264488, 264259, 264508, 264905, 264906, 264907, 264908, 264909, 264510, 264511, 264512, 265008, 265009, 264910, 264591, 264596, 264759, 265010, 265011, 18108351, 264763, 264288, 264768, 264788, 264693, 18108370, 264629, 18108372, 264830, 264631, 264834, 264558, 18108385, 264482, 264584, 264567
744	80230421 (1487, 1488)				18108397, 264511, 264690, 264628, 264636, 264692, 264639, 264768
745	9841863 (1489, 1490)	Novel Protein sim. GBank gi 78921 pri J504846 - UDP-N- acetyluramoylalanine-D-glutamate-2, 8-diaminopimelate-D- alanine-D-alanine ligase (EC 6.3.2.15) precursor - Escherichia coli	glycoprotein		264906
746	11073229 (1491, 1492)	Novel Protein sim. GBank gi 3366354 (AF074705) - pyocitellin synthetase [Pseudomonas aeruginosa]	synthase		264600
747	94322044 (1493, 1494)	Novel Protein sim. GBank gi 286741 dbj BA24848 - (AB007678) KIAA0418 [Homo sapiens]	oxidase	Contains protein domain (PF00018) - SH3 domain	66714117, 264905, 264509, 264906, 264907, 264908, 264909, 264511, 264910, 265011, 264881, 264288, 264766, 264887, 264768, 264769, 21906768, 35895917, 264891, 264693, 264828, 264834, 264835, 264839, 56182323, 83373044
748	11617923 (1495, 1498)				264690

749	20469119 (1497, 1498)	Novel Protein sim. GBank gi11169727 sp P44948 PPG_HAEIN - FORMAMIDOPYRIMIDINE-DNA GLYCOSYLASE (FAPY- DNA GLYCOSYLASE)		UNCLASSIFIED	264804 264600
750	20298427 (1499, 1500)				
751	21635169 (1501, 1502)	Novel Protein sim. GBank gi15360066 gb AAD2851.1 AF15988 - (AF15988) serine/threonine kinase PKN3 [Myxococcus xanthus]	Contains protein domain (PF00069) - kinase Eukaryotic protein kinase domain		264605, 264559
752	82450366 (1503, 1504)	Novel Protein sim. GBank gi1168882 sp P44426 BIOA_HAEIN - ADENOSYLMETHIONINE-8-AMINO-7-OXONONANOATE AMINOTRANSFERASE (7,8-DIAMINO-PELAGONIC ACID AMINOTRANSFERASE) (DAPA AMINOTRANSFERASE)	Contains protein domain (PF00202) - Aminotransferases class-III pyridoxal phosphate	UNCLASSIFIED	264506, 264907, 284510, 265011, 264782, 284889, 35895855, 284638, 18106387
753	80506718 (1505, 1506)	Novel Protein sim. GBank gi12651530 sp P32399 YHGE_BACSU - HYPOTHETICAL 84.1 KD PROTEIN IN HEMY-GLTT INTERGENIC REGION (ORF8)		UNCLASSIFIED	284908, 284800, 284602, 264604, 284780, 284769, 264634
754	85083741 (1507, 1508)				
755	80185449 (1509, 1510)			UNCLASSIFIED	264506, 264906, 284907, 284908, 284909, 284759, 284802, 284784, 284769, 284628, 284629, 284630, 264632, 264634, 264635, 264637, 284638, 63373044, 18108365
756	94631886 (1511, 1512)	Novel Protein sim. GBank gi13449276 emb CAA20420 - (AL031317) putative dehydrogenase [Streptomyces coelicolor]		UNCLASSIFIED	264448, 264690
757	78486533 (1513, 1514)				284769, 264689, 264638, 264639
758	78663176 (1515, 1518)	Novel Protein sim. GBank gi14580331 emb CAB40107.1 - (AJ001206) putative glycogen debranching enzyme [Streptomyces coelicolor]		UNCLASSIFIED	264682, 264665
759	78475667 (1517, 1518)	Novel Protein sim. GBank gi12911858 (AF047659) - No definition line found [Caenorhabditis elegans]	amylase		285007, 16108387, 265007, 18108387
760	87628888 (1519, 1520)	Novel Protein sim. GBank gi13451312 emb CAA20449 - (AL031324) membrane alipase [Schizosaccharomyces pombe]	Contains protein domain (PF00122) - E1-E2 ATPase	UNCLASSIFIED	264684, 264686
761	79877988 (1521, 1522)	Novel Protein sim. GBank gi13327156 dbj BAA31647 - (AB014572) KIAA0872 protein [Homo sapiens]			29331822, 264908, 52644045, 56162435, 60170831, 21906754, 265017, 265019, 264681, 284687, 264688, 21906766, 21908768, 265020, 265021, 265022, 264835, 22278000
762	80023583 (1523, 1524)			UNCLASSIFIED	264786
763	20204813 (1525, 1526)	Novel Protein sim. GBank gi14981266 gb AAD35822.1 AE00174 - (AE00174) lipopolysaccharide core biosynthesis protein KdIB [Thermotoga maritima]		UNCLASSIFIED	284907, 264593, 265020
764	39515024 (1527, 1528)				264600
					264603

765	80025347 (1528, 1530)	Novel Protein sim. GBank gij3845083 (AE001371) - erythrocyte membrane protein PIEMP3 [Plasmodium falciparum]		struct	264905, 264906, 264594, 264888, 33857023
766	82417404 (1531, 1532)			UNCLASSIFIED	264805, 264762, 18108374
767	10298742 (1533, 1534)	Novel Protein sim. GBank gij54112 [pir]S40827 - hypothetical protein 6300 - Escherichia coli		UNCLASSIFIED	264907
768	79416080 (1535, 1536)	Novel Protein sim. GBank gij2882501 [emb]CAA06164 - (AJ004632) neuropathy target esterase [Homo sapiens]		UNCLASSIFIED	264592, 264595
769	80086554 (1537, 1538)	Novel Protein sim. GBank gij283437 [pir]S27850 - hypothetical protein - Trypanosoma cruzi (fragment)		UNCLASSIFIED	55810764, 264559
770	80417847 (1539, 1540)			UNCLASSIFIED	264905, 264907, 264826, 264908, 265010, 264766, 264628, 264629, 264634, 264638, 284555
771	95328508 (1541, 1542)	Novel Protein sim. GBank gij4788004 [gb]AAD29715.1 [AF140598] - (AF140598) ring-box protein 1 [Homo sapiens]	Contains protein domain (PF00097) - Zinc finger, C3HC4 type (RING finger)	UNCLASSIFIED	56182575, 35696286, 264259, 66714117, 264107, 66712502, 56182435, 284112, 55612038, 87168559, 264268, 21908766, 33857023, 65274620, 65274791, 18108381
772	78971362 (1543, 1544)			UNCLASSIFIED	264910
773	78945363 (1545, 1546)	Novel Protein sim. GBank gij5531324 [emb]CAB51045.1 - (AJ009579) putative alkane 1-monooxygenase [Pseudomonas fluorescens]		UNCLASSIFIED	265020
774	79856128 (1547, 1548)			UNCLASSIFIED	264909
775	20620141 (1549, 1550)			UNCLASSIFIED	264555
776	78942693 (1551, 1552)		Contains protein domain (PF01008) - Hepatitis C virus non-structural protein NS4a		285019
777	79960376 (1553, 1554)	Novel Protein sim. GBank gij4505481 [ref]NP_003624.1 [pNRPB - nuclear restricted protein, BTB domain-like (brain)]	Contains protein domain (PF01344) - Kelch motif	protease	21908754, 265020, 60170615, 264691
778	20891310 (1555, 1556)			UNCLASSIFIED	264511
779	80054024 (1557, 1558)				264803
780	95268987 (1559, 1560)	Novel Protein sim. GBank gij1144520 (U34956) - phosphoribosylformylglycinamide synthase [Mycobacterium tuberculosis]		synthase	264907, 264600, 264601, 264602, 264603, 264604, 264605, 264486
781	80250049 (1561, 1562)				264905, 264907, 285010, 264600, 264601, 18108362, 18108374, 264556
782	8758529 (1563, 1564)	Novel Protein sim. GBank gij4155447 (AE001517) - proline/betaine transporter [Helicobacter pylori J99]		UNCLASSIFIED	264605
783	18410791 (1565, 1566)				285020
784	80051197 (1587, 1588)			UNCLASSIFIED	264635, 33657023, 29331828, 265017, 264555, 264568
785	56073541 (1569, 1570)	Novel Protein sim. GBank gij3451335 (AC005525) - F22162_1 [Homo sapiens]	Contains protein domain (PF00047) - immunoglobulin domain	struct	35698052, 284804
786	20438842 (1571, 1572)	Novel Protein sim. GBank gij138748 [sp]P10905 [UGPA, ECOLI - SN-GLYCEROL-3-PHOSPHATE TRANSPORT SYSTEM PERMEASE PROTEIN UGPA]		transport	264603
787	80258364 (1573, 1574)			UNCLASSIFIED	264593

788	80507844 (1575, 1576)	Novel Protein sim. GBank gll2748078 (AF015310) - BTH1 (Brassica napus)		synthese	284808, 284802, 284803, 284769, 284638
789	17294715 (1577, 1578)	Novel Protein sim. GBank gll2351849 (U83357) - 40 kDa heat shock chaperone protein (Halobacterium salinarum)		UNCLASSIFIED	285007
790	86264408 (1579, 1580)	Novel Protein sim. GBank gll5706378 (dbj BAA03099.1) - (AB028118) MALT1 [Homo sapiens]	Contains protein domain (PF00047) - Immunoglobulin domain	glycoprotein	22278989, 29331824, 284828, 87188559, 265018, 21906765, 21906787, 21908766, 21905769, 265020, 284692, 22278000, 284583
791	94651827 (1581, 1582)	Novel Protein sim. GBank gll5689948 (emb CAB51885.1) - (AL109863) putative isoleucyl-tRNA synthetase (Streptomyces coelicolor A3(2))			284801, 284605, 284638
792	80058788 (1583, 1584)	Novel Protein sim. GBank gll393194 (L02375) - S-antigen (Plasmodium falciparum)		struct	285021, 284631, 284635, 284556
793	79838730 (1585, 1586)	Novel Protein sim. GBank gll1345408 (dbj BAA05046) - (D26046) AT motif-binding factor [Mus musculus]	Contains protein domain (PF00046) - Homeobox domain	homeobox	264893
794	81839284 (1587, 1588)	Novel Protein sim. GBank gll105884 (pir S24023) - dopamine receptor D4 - human (fragment)		UNCLASSIFIED	264603, 284604, 264910, 284762, 264808, 284838, 284908, 284757
795	80074988 (1589, 1590)	Novel Protein sim. GBank gll1877334 (emb CAB07082) - (Z92771) bfrA [Mycobacterium tuberculosis]		carboxylase	284488, 35686052, 284805, 264807, 265010, 35698423, 284636
796	86669451 (1591, 1592)	Novel Protein sim. GBank gll1877334 (emb CAB07082) - (Z92771) bfrA [Mycobacterium tuberculosis]			60432228, 55811150, 284630, 284637, 284585
797	87771781 (1593, 1594)	Novel Protein sim. GBank gll2985447 (emb CAA71518) - (Y10495) CDV-1R protein [Mus musculus]		struct	22278988, 264083, 264084, 66714117, 21908787, 21906769, 265020, 285022
798	79865209 (1595, 1596)	Novel Protein sim. GBank gll4467250 (emb CAB37575) - (AL035589) probable Glu-tRNA Gln amidotransferase subunit [Streptomyces coelicolor]		transcript factor	284687, 264768, 284683
799	79557818 (1597, 1598)	Novel Protein sim. GBank gll4467250 (emb CAB37575) - (AL035589) probable Glu-tRNA Gln amidotransferase subunit [Streptomyces coelicolor]		hydrolase	264908, 264910, 264638, 264638
800	79970188 (1599, 1600)	Novel Protein sim. GBank gll2791517 (emb CAA16054) - (AL021246) hypothetical protein Rv2477c [Mycobacterium tuberculosis]		UNCLASSIFIED	284488
801	80489399 (1601, 1602)	Novel Protein sim. GBank gll4887211 (gb AAD32237.1) AF14744 - (AF147449) penicillin binding protein 1B [Pseudomonas aeruginosa]		transport	284508, 284511, 265008, 265009, 264769, 264587, 284486
802	79834598 (1603, 1604)	Novel Protein sim. GBank gll4887211 (gb AAD32237.1) AF14744 - (AF147449) penicillin binding protein 1B [Pseudomonas aeruginosa]			264905, 284693
803	20467520 (1605, 1606)	Novel Protein sim. GBank gll1178152 (gb P44507) YHAD_HAEIN - HYPOTHETICAL PROTEIN H10091		struct	264605
804	10174239 (1607, 1608)	Novel Protein sim. GBank gll1178152 (gb P44507) YHAD_HAEIN - HYPOTHETICAL PROTEIN H10091		kinase	264510
805	78599993 (1609, 1610)	Novel Protein sim. GBank gll2764612 (emb CAA04683) - (AJ001330) ornithine transcarbamoylase [Lactobacillus sakei]	Contains protein domain (PF00185) - Aspartate/ornithine carbamoyltransferase		284508
806	80464113 (1611, 1612)	Novel Protein sim. GBank gll2764612 (emb CAA04683) - (AJ001330) ornithine transcarbamoylase [Lactobacillus sakei]		transferase	284769
807	80381812 (1613, 1614)	Novel Protein sim. GBank gll2833311 (sp Q21829) YNFD_CAEEL - HYPOTHETICAL 18.9 KD PROTEIN R07E5.13 IN CHROMOSOME III			264764

806	35108817 (1615, 1616)	Novel Protein sim. GBank gi 3913092 sp Q46170 ARCD_CLOPE - ARGININE/ORNITHINE ANTIPORTER		transport	264909, 264602, 21908764, 18108374
809	81454254 (1617, 1618)	Novel Protein sim. GBank gi 3913019 sp P74309 ALF1_SYNY3 - FRUCTOSE- BISPHOSPHATE ALDOOLASE CLASS I (FBP ALDOOLASE)	Contains protein domain (PF00274) - Fructose-bisphosphate aldolase class-I	UNCLASSIFIED	264508, 264906, 264909, 265007, 284910, 264758, 284800, 264602, 264603, 264805, 284687, 284769, 284889, 284636, 284486 284389
810	80192781 (1619, 1620)	Novel Protein sim. GBank gi 401472 sp P30863 YAFB_ECOLI - HYPOTHETICAL OXIDOREDUCTASE IN ASPU-MLTD INTERGENIC REGION	Contains protein domain (PF00248) - Aldoketo reductase family	reductase	
811	80079280 (1621, 1622)			UNCLASSIFIED	284556
812	10297654 (1623, 1624)			UNCLASSIFIED	284692
813	78612280 (1625, 1626)			UNCLASSIFIED	284908
814	80473427 (1627, 1628)	Novel Protein sim. GBank gi 146168 (J01617) - glutaminyl- [RNA synthetase [Escherichia coli]		synthase	284905, 264602, 284605, 264682, 264887, 284769, 284636
815	95419513 (1629, 1630)	Novel Protein sim. GBank gi 4589652 dbj BAA76848.1 - (AB023221) KIAA1004 protein [Homo sapiens]		UNCLASSIFIED	284488, 22278998, 22278999, 29331822, 29331624, 29331825, 29331827, 29331828, 29146499, 284905, 284908, 265007, 33657402, 60433356, 60433438, 264758, 285011, 285017, 285018, 285019, 284369, 264288, 284885, 21906765, 21906767, 285020, 285021, 284892, 85274820, 33657109, 284828, 18108376, 284635, 264838, 60170394, 58162323, 284564 284600
816	19881910 (1631, 1632)	Novel Protein sim. GBank gi 1781144 emb CA806254 - (Z63866) hypothetical protein Rv3069 [Mycobacterium tuberculosis]		UNCLASSIFIED	264595
817	95293316 (1633, 1634)	Novel Protein sim. GBank gi 1781144 emb CA806254 - (Z63866) hypothetical protein Rv3069 [Mycobacterium tuberculosis]		UNCLASSIFIED	
818	80938190 (1635, 1636)	Novel Protein sim. GBank gi 1477468 (U35244) - vacuolar protein sorting homolog r-vps33a [Rattus norvegicus]			85274572, 22278999, 80424269, 35696052, 55812036, 21906766, 55811957, 35695917, 33657023, 18108370, 18108374, 55810764, 35696423, 55811576, 264636 264565
819	80254977 (1637, 1638)	Novel Protein sim. GBank gi 1001352 dbj BAA10839 - (D64008) ABC transporter [Synecocystis sp.]		transport	
820	60059688 (1639, 1640)	Novel Protein sim. GBank gi 586814 sp P37484 YYBT_BACSU - HYPOTHETICAL 74.3 KD PROTEIN IN RPLI-COTF INTERGENIC REGION		UNCLASSIFIED	284600, 284802, 284804
821	79762590 (1641, 1642)			UNCLASSIFIED	264910
822	80215310 (1643, 1644)			UNCLASSIFIED	284510, 284594, 284637
823	94992299 (1645, 1646)	Novel Protein sim. GBank gi 3878400 emb CAA95628 - (Z71264) predicted using GeneFinder. Weak similarity to Mouse T-complex-associated-testes-expressed-1 protein (PIR Acc. No. A45841); cDNA EST EMBL:D32742 comes from this gene; cDNA EST EMBL:D33817 comes from this gene; cDNA EST...		struct	264509, 284687, 264691
824	80411171 (1647, 1648)	Novel Protein sim. GBank gi 1370076 emb CAA66887 - (X98235) type I [Drosophila melanogaster]	Contains protein domain (PF01429) - Methyl-CpG binding domain		284910, 284783, 264769, 284693

825	20638600 (1648, 1650)	Novel Protein sim. GBank gi 3025132 sp P77391 YEAG_ECOLI - HYPOTHETICAL 74.5 KD. PROTEIN IN GAP-RND INTERGENIC REGION		UNCLASSIFIED	264592
826	11075047 (1651, 1652)	Novel Protein sim. GBank gi 3242281 emb CAA16668 - (AL021848) hypothetical protein Rv3202c [Mycobacterium tuberculosis]			264605
827	80064207 (1653, 1654)	Novel Protein sim. GBank gi 3417424 emb CAA20312 - (AL031261) putative transport protein [Schizosaccharomyces pombe]			264603
828	95106322 (1655, 1656)	Novel Protein sim. GBank gi 4336882 pb AAD17887 - (AF101361) Abnormal X segregation [Drosophila melanogaster]		UNCLASSIFIED	52845080, 264508, 264905, 284509, 284906, 284807, 264906, 264908, 264512, 264910, 284591, 264758, 264800, 264766, 264768, 21906768, 35695917, 284691, 264628, 264630, 284631, 264632, 284634, 264635, 264636, 264637, 264558, 284639, 83373044, 18108385, 264583, 264566, 284488
829	81742215 (1657, 1658)	Novel Protein sim. GBank gi 3820539 (AF080002) - UDP-N- acetylmuramyl tripeptide synthetase MurC [Helicobacter mobilis]		UNCLASSIFIED	264758, 264634
830	20396091 (1659, 1660)			UNCLASSIFIED	264603
831	67112435 (1661, 1662)	Novel Protein sim. GBank gi 1870004 emb CAB06655 - (Z92539) hypothetical protein Rv1024 [Mycobacterium tuberculosis]		UNCLASSIFIED	66714117, 284910, 264639
832	19536322 (1663, 1664)			UNCLASSIFIED	284906
833	20726654 (1665, 1666)	Novel Protein sim. GBank gi 2500058 sp Q46267 PFLA_CLOPA - PYRUVATE FORMATE-LYASE ACTIVATING ENZYME		UNCLASSIFIED	264602
834	21428762 (1667, 1668)	Novel Protein sim. GBank gi 2497531 sp Q46078 KPYK_CORGL - PYRUVATE KINASE (PK)	Contains protein domain (PF00224) - Pyruvate kinase		284600, 264602, 264769, 264689, 264636
835	94140482 (1669, 1670)				264768, 263994, 21906767, 284910, 264632, 284635, 284259, 264639, 264893, 83373044, 284758, 35696052, 22279002, 264508, 284905, 264806, 264448, 283972, 264908, 284909
836	66126552 (1671, 1672)	Novel Protein sim. GBank gi 699315 (U15184) - phosphate transport protein PSTC [Mycobacterium leprae]	transport		35695917, 264557
837	79450450 (1673, 1674)	Novel Protein sim. GBank gi 728867 sp P40602 APG_ARATH - ANTER-SPECIFIC PROLINE-RICH PROTEIN APG PRECURSOR		UNCLASSIFIED	264595
838	79184203 (1675, 1676)			UNCLASSIFIED	264687
839	78641125 (1677, 1678)	Novel Protein sim. GBank gi 2498533 sp Q50598 YOD8_MYCTU - HYPOTHETICAL 69.9 KD PROTEIN CY1A11.08		UNCLASSIFIED	264906
840	80058051 (1679, 1680)	Novel Protein sim. GBank gi 4557753 ref NP_000372.1 pMID1 - midline 1 protein (finger)	Contains protein domain (PF00097) - Zinc finger, C3HC4 type (RING finger)		264762, 284556

841	80378318 (1681, 1882)	Novel Protein sim. GBank gl139805[sp]P08045XFEIN_XENLA - XFEIN PROTEIN	Contains protein domain (PF00096) - Zinc finger, C2H2 type	transcription factor	264784
842	80078724 (1683, 1884)	Novel Protein sim. GBank gl12114321[db][BAA200371] - (D88733) membrane glycoprotein [Equine herpesvirus 1]	Contains protein domain (PF00569) - Zinc finger present in dystrophin, CBP/p300	UNCLASSIFIED	264905, 284906, 265008, 265008, 16108374, 56182323, 284558
843	67002847 (1665, 1686)	Novel Protein sim. GBank gl13882325[db][BAA34522.1] - (AB016345) KIAA0802 protein [Homo sapiens]	Contains protein domain (PF00170) - bZIP transcription factor	struc	264091, 28331625, 284806, 284768, 284563
844	17841439 (1667, 1686)	Novel Protein sim. GBank gl12224721[db][BAA20844] - (AB002386) KIAA0380 [Homo sapiens]	Contains protein domain (PF00086) - Zinc finger, C2H2 type	transcription factor	265011
845	18346844 (1689, 1680)	Novel Protein sim. GBank gl1625879[pi][A36828] - virulence regulatory protein VsrB - Pseudomonas solanacearum		kinase	264629
846	78863441 (1691, 1692)				284907
847	76695346 (1693, 1694)				284909
848	78489385 (1695, 1686)			UNCLASSIFIED	265020
849	78758367 (1687, 1686)			UNCLASSIFIED	264568
850	76817649 (1699, 1700)	Novel Protein sim. GBank gl13183245[sp]P78061[CYJK_ECOLI] - PUTATIVE GLUTAMINE SYNTHETASE (GLUTAMATE-AMMONIA LIGASE)	Contains protein domain (PF00120) - Glutamine synthetase	UNCLASSIFIED	284909
851	95320333 (1701, 1702)	Novel Protein sim. GBank gl15454130[ref][NP_006280.1p.TLN] - talin	Contains protein domain (PF01808) - ILWEEQ domain		264488, 52644507, 264488, 16108388, 65274572, 56162575, 22278994, 22278995, 22278998, 35898288, 22278997, 22278998, 22278998, 20261171, 264480, 264259, 52845080, 29331822, 29331624, 66714117, 29331825, 60432289, 28331628, 28331827, 35696052, 28331828, 29146488, 29146499, 264107, 264905, 264906, 264907, 264908, 52644045, 58182435, 265008, 265007, 285008, 285008, 284910, 60432228, 60431735, 60433358, 33857402, 80433438, 264595, 264758, 264759, 21908754, 33109854, 52644286, 265010, 265011, 87168559, 285017, 285018, 285019, 284760, 264761, 264762, 264681, 16108351, 264763, 264446, 264882, 264764, 264883, 16108354, 264288, 264369, 284685, 264766, 264687, 264766, 264769, 21906765, 21906766, 21906767, 21906768, 29148627, 21906769, 29148628, 55811957, 35695917, 265020, 285021, 285022, 60170615, 52644150, 264691, 264692, 33857023, 264693, 263966, 33657108, 27486281, 27486282, 27486264, 27486265, 35685783, 60431602, 16108370, 20281088, 284828, 16108374, 16108376, 55811578, 35898423, 35895855, 284634, 264635, 264636, 264555, 60431650, 264556, 264691
852	10147388 (1703, 1704)				

853	13032587 (1705, 1708)	Novel Protein sim. GBank gi 3402836 emb CAA780821 - (Y16136) 2-oxoate reductase [Moorella thermoacetica]		reductase	264638 264566
854	80052438 (1707, 1708)				
855	79641130 (1708, 1710)				264692
856	11594238 (1711, 1712)			UNCLASSIFIED	264591
857	79210165 (1713, 1714)			UNCLASSIFIED	264630, 264634
858	80248910 (1715, 1716)				265008, 265009, 264601, 264602, 264603, 18108351
859	20296634 (1717, 1718)				264559
860	80041749 (1719, 1720)			UNCLASSIFIED	264489
861	65857045 (1721, 1722)			UNCLASSIFIED	33657023, 264630
862	80079487 (1723, 1724)			UNCLASSIFIED	264600
863	80578831 (1725, 1726)	Novel Protein sim. GBank gi 2248532 (U83872) - ORF 73, contains large complex repeat CR 73 [Kaposi's sarcoma-associated herpesvirus]		UNCLASSIFIED	264488, 18108388, 35696286, 264259, 16106351, 264288, 265021
864	94939904 (1727, 1728)				
865	80045310 (1729, 1730)	Novel Protein sim. GBank gi 5889884 emb CA852047.1 - (AL109732) hypothetical protein [Streptomyces coelicolor A3(2)]	Contains protein domain (PF01479) - S4 domain	UNCLASSIFIED	264259, 264112, 263974
866	80162031 (1731, 1732)	Novel Protein sim. GBank gi 4557870 ref NP_000341.1 pABCR - ATP binding cassette transporter		transport	264635, 264600, 264638, 264581, 264602, 264693
867	80062402 (1733, 1734)				284288, 284557, 264558
868	10075384 (1735, 1736)				284605
869	80062406 (1737, 1738)			UNCLASSIFIED	284909
870	80249651 (1739, 1740)	Novel Protein sim. GBank gi 628660 pir S37755 - Adenyltransferase - Escherichia coli		transferase	264605, 284887, 18108374
871	20378295 (1741, 1742)	Novel Protein sim. GBank gi 1708180 ref Q10602 HEMK_MYCTU - HEMK PROTEIN HOMOLOG		UNCLASSIFIED	264601, 284636
872	95187114 (1743, 1744)	Novel Protein sim. GBank gi 1545959 emb CAA677631 - (X99384) paladin [Mus musculus]			284603
				UNCLASSIFIED	35696288, 22278988, 264259, 29331822, 29331824, 29331825, 60432289, 29331826, 29331827, 29331828, 35896052, 264509, 284905, 284906, 284907, 284908, 284909, 264510, 285008, 284511, 284512, 285007, 285008, 285009, 284910, 284591, 60433356, 284596, 52848317, 87168474, 265010, 284602, 284803, 265017, 265018, 284605, 18108351, 264764, 264766, 264768, 52844228, 284789, 21906765, 265021, 264534, 284691, 52845128, 264628, 284629, 35698423, 65274791, 284631, 284632, 264635, 264636, 284556, 264637, 284638, 264839, 60432113, 22279000, 22279002, 284584
673	20189728 (1745, 1746)	Novel Protein sim. GBank gi 4158104 (AE001589) - putative Outer membrane protein [Helicobacter pylori J98]	Contains protein domain (PF00098) - Zinc finger, C2H2 type	UNCLASSIFIED	284595

874	80077602 (1747, 1748)	Novel Protein sim. GBank gi134318 sp P07818 ISCRB_BACSU - SUCROSE-6- PHOSPHATE HYDROLASE (SUCRASE) (INVERTASE)		UNCLASSIFIED	284600
875	86600448 (1749, 1750)	Novel Protein sim. GBank gi1481000 p S37594 - muchin - human (fragment)			284259, 284448, 284288, 284557, 87168518
876	86465157 (1751, 1752)	Novel Protein sim. GBank gi13128283 (AF010495) - iron(III) diclrate transport ATP-binding protein [Rhodobacter capsulatus]	Contains protein domain (PF00005) - ABC transporter	transport	284907, 284601, 284602, 284605, 285020, 80431802
877	87802548 (1753, 1754)	Novel Protein sim. GBank gi1731074 sp P40349 URB1_USTMA - SIDEROPHORE BIOSYNTHESIS REGULATORY PROTEIN URBS1	Contains protein domain (PF00320) - GATA zinc finger	transcriptionfactor	22278998, 284909, 284389
878	80187289 (1755, 1756)	Novel Protein sim. GBank gi1351614 sp Q09853 YAEED_SCHPO - HYPOTHETICAL 181.5 KD PROTEIN C23D3.13C IN CHROMOSOME I		ATPase_associated	284389, 284555
879	94328962 (1757, 1758)	Novel Protein sim. GBank gi13875304 emb CAA84341 - (Z74030) predicted using GeneFinder; cDNA EST EMBL: C07808 comes from this gene; cDNA EST EMBL: C08023 comes from this gene; cDNA EST yk4505e9.3 comes from this gene; cDNA EST yk489hg.3 comes from this gene; cDNA EST yk489hg.5 com...			56182575, 28331824, 284508, 284008, 285018, 18108351, 284448, 284883, 21808768, 21908768, 80170815, 33657023, 65274820, 33857109, 18108374, 35895855, 284583
880	6491135 (1759, 1760)	Novel Protein sim. GBank gi137120 sp P11214 UROT_MOUSE - TISSUE PLASMINOGEN ACTIVATOR PRECURSOR (TPA) (T- PLASMINOGEN ACTIVATOR)	Contains protein domain (PF00051) - Kringie domain	calhepsin	284508
881	11280122 (1761, 1762)				
882	11077011 (1763, 1764)	Novel Protein sim. GBank gi1263209 emb CAA75697 - (Y15513) Prodes protein [Drosophila melanogaster]		UNCLASSIFIED	264508
883	79582968 (1765, 1766)				284558
884	13517921 (1767, 1768)	Novel Protein sim. GBank gi115506 emb CAA64425 - (X94976) cell wall-plasma membrane linker protein [Brassica napus]		UNCLASSIFIED	264688
885	80052457 (1769, 1770)	Novel Protein sim. GBank gi12078027 emb CAB084671 - (Z95208) hypothetical protein Rv2372c [Mycobacterium tuberculosis]		UNCLASSIFIED	264638
886	11885138 (1771, 1772)				
887	94315307 (1773, 1774)	Novel Protein sim. GBank gi12695834 emb CAA15904 - (AL021009) aucA [Mycobacterium tuberculosis]			284690
888	10083399 (1775, 1776)	Novel Protein sim. GBank gi15689395 emb BAA82981.1 - (AB028952) KIAA1029 protein [Homo sapiens]		dehydrogenase	35698052, 264908, 264600, 264603, 35895917, 35695855, 284838
889	20385917 (1777, 1778)	Novel Protein sim. GBank gi11881338 emb BAA18365 - (AB001488) PROBABLE INTEGRAL MEMBRANE PROTEIN, SIMILAR TO CHLORAMPHENICOL RESISTANCE PROTEIN OF STREPTOMYCES VENEZUELAE [Bacillus subtilis]		UNCLASSIFIED	284808
890	18904337 (1779, 1780)	Novel Protein sim. GBank gi1854085 emb CAA58337 - (X83413) U88 [Human herpesvirus 6]			284603

891	13518879 (1761, 1782)	Novel Protein sim. GBank gi4959398 gb AA034331.1 AF11248 - (AF112481) RAD54B Protein [Homo sapiens]		UNCLASSIFIED	264638
892	87634157 (1763, 1784)	Novel Protein sim. GBank gi454526 gb 143833 - LBP- 1b-transcription factor binding to initiation site of HIV-1 (alternatively spliced) [human, Namalwa cells, Peptide, 541 aa]		transcription factor	22278996, 22278999, 29331828, 35696052, 264908, 264909, 285009, 285011, 264802, 265019, 264766, 21906765, 21906768, 21908769, 285020, 285021, 56528488 264889, 263987
893	79166037 (1765, 1788)	Novel Protein sim. GBank gi2829888 sp P80608 CYSK_MAIZE - CYSTEINE SYNTHASE (O-ACETYL SERINE SULFHYDRYLASE) (O- ACETYL SERINE (THIOL)-LYASE) (CSASE)		synthase	
894	11102240 (1787, 1788)				263978
895	80239868 (1789, 1790)			UNCLASSIFIED	264508, 264600, 264555, 264559
896	79747803 (1791, 1792)				264632
897	94991923 (1793, 1794)			UNCLASSIFIED	264888, 28331828, 264511
898	87895109 (1795, 1796)		Contains protein domain (PF00047) - Immunoglobulin domain		
899	11100463 (1797, 1798)				56182575, 60432289, 56182435, 60432229, 55811957, 22279000, 264488
900	80499768 (1799, 1800)				264801
901	80502410 (1801, 1802)	Novel Protein sim. GBank gi1750127 (U66480) - YncC [Bacillus subtilis]	transport		264789, 264891, 264563
902	80503301 (1803, 1804)	Novel Protein sim. GBank gi312287 sp O07438 SVA_MYCTU - ALANYL-TRNA SYNTHETASE (ALANINE-TRNA LIGASE) (ALARS)	Contains protein domain (PF01411) - RNA synthetases class II (A)	synthase	264907, 264602, 264805, 264789, 35895917, 18108376, 264563
903	82080208 (1805, 1806)	Novel Protein sim. GBank gi335570 fmb CAA200011 - (AL031124) 3-isopropylmalate dehydratase large subunit [Streptomyces coelicolor]	Contains protein domain (PF00330) - Aconitase family (aconitase hydratase)	isomerase	264909, 285008, 264802, 264804, 264789, 264889, 264693
904	20451078 (1807, 1808)	Novel Protein sim. GBank gi2960120 fmb CAA18018.1 - (AL022121) glpK [Mycobacterium tuberculosis]	Contains protein domain (PF00370) - FGCY family of carbohydrate kinases	kinase	35698052, 284905, 264510, 264511, 264512, 264605, 264760, 18108351, 264782, 264887, 264786, 264769, 264888, 21908764, 35695917, 27486282, 35895855, 264634, 264636, 264488
905	9398463 (1809, 1810)	Novel Protein sim. GBank gi728887 sp P40908 ARGI COCIM - ARGINASE	Contains protein domain (PF00491) - Arginase family	hydrolyase	264604
906	80052626 (1811, 1812)	Novel Protein sim. GBank gi4567200 gb AA023616.1 AC00718 - (AC007186) hypothetical protein [Arabidopsis thaliana]	UNCLASSIFIED		264909
907	87813201 (1813, 1814)	Novel Protein sim. GBank gi3368894 gb BAA342961 - (AB015054) Alg2 [Rhizomucor pusillus]			284595, 284605
908	11754482 (1815, 1816)	Novel Protein sim. GBank gi4569729 gb BAA78863.1 - (AB003137) DnaJ homolog protein [Saltia gligiana]	UNCLASSIFIED		60432289, 264601, 264680
909	20727907 (1817, 1818)			UNCLASSIFIED	284636
910	16776206 (1819, 1820)		Contains protein domain (PF00684) - DnaJ central domain (4 repeats)	eph	264602 265009

911	87454340 (1821, 1822)	Novel Protein sim. GBank gi 546774 sp P35885 RL7A_ORYSA - 60S RIBOSOMAL PROTEIN L7A		ribosomal prot	285010, 264604, 80432113
912	20448663 (1823, 1824)	Novel Protein sim. GBank gi 2314008 gb AAD07921.1 - (AE000597) CDP-diglyceride hydrolase (cdh) (Helicobacter pylori 26895)		hydrolase	264559
913	20469357 (1825, 1826)	Novel Protein sim. GBank gi 17857 sp Q03804 RIR1_CAEEL - PROBABLE RIBONUCLEOSIDE-DIPHOSPHATE REDUCTASE LARGE CHAIN (RIBONUCLEOTIDE REDUCTASE)		UNCLASSIFIED	264604
914	79183351 (1827, 1828)	Novel Protein sim. GBank gi 5689571 db BAA83069.1 - (AB029040) KIAA1117 protein (Homo sapiens)	Contains protein domain (PF00317) - Ribonucleotide reductase	reductase	284636
915	87606703 (1829, 1830)	Novel Protein sim. GBank gi 186110 emb CAA71790 - (Y10831) putative integrase (Rickettsia eutropha)		UNCLASSIFIED	18108398, 22278998, 66714117, 264608, 264591, 21908788, 285020, 55811578, 264638
916	79444091 (1831, 1832)	Novel Protein sim. GBank gi 1655899 emb CAA69032 - (Y07752) pterophorin-S (Volvox carter)		UNCLASSIFIED	264595
917	20195985 (1833, 1834)	Novel Protein sim. GBank gi 5689968 emb CAB52005.1 - (AL108663) putative membrane protein (Streptomyces coelicolor A3(2))		synthase	264605
918	91228785 (1835, 1836)	Novel Protein sim. GBank gi 168448 sp Q05813 AMP1_STRLI - XAA-PRO AMINOPEPTIDASE I (X-PRO AMINOPEPTIDASE I) (AMINOPEPTIDASE P I) (APP) (PEPP I)			264299, 29331826, 264908, 265018, 264448, 265020, 284635, 83373044
919	80436785 (1837, 1838)	Novel Protein sim. GBank gi 3650084 emb CAA21911.1 - (AL033389) alcohol dehydrogenase (Schizosaccharomyces pombe)			265006, 284512, 264600, 284602, 264604, 264788, 18108370, 284583
920	79698095 (1839, 1840)	Novel Protein sim. GBank gi 12497888 sp Q06983 PAFA_MOUSE - PLATELET- ACTIVATING FACTOR ACETYLHYDROLASE PRECURSOR (PAF ACETYLHYDROLASE) (PAF 2- ACYLHYDROLASE) (LDL-ASSOCIATED PHOSPHOLIPASE A2) (LDL-PLA2) (2-ACETYL-1- ALKYLGLYCEROPHOSPHOCHOLINE ESTERASE) (1- ALKYL-2-ACETYLGLYCEROPHO...		peptidase	264508
921	18658634 (1841, 1842)	Novel Protein sim. GBank gi 287079 sp P29514 TB86_ARATH - TUBULIN BETA-8 CHAIN		UNCLASSIFIED	264600
922	78982605 (1843, 1844)	Novel Protein sim. GBank gi 2497888 sp Q06983 PAFA_MOUSE - PLATELET- ACTIVATING FACTOR ACETYLHYDROLASE PRECURSOR (PAF ACETYLHYDROLASE) (PAF 2- ACYLHYDROLASE) (LDL-ASSOCIATED PHOSPHOLIPASE A2) (LDL-PLA2) (2-ACETYL-1- ALKYLGLYCEROPHOSPHOCHOLINE ESTERASE) (1- ALKYL-2-ACETYLGLYCEROPHO...		tubulin	285019, 22279002
923	86895630 (1845, 1846)	Novel Protein sim. GBank gi 2497888 sp Q06983 PAFA_MOUSE - PLATELET- ACTIVATING FACTOR ACETYLHYDROLASE PRECURSOR (PAF ACETYLHYDROLASE) (PAF 2- ACYLHYDROLASE) (LDL-ASSOCIATED PHOSPHOLIPASE A2) (LDL-PLA2) (2-ACETYL-1- ALKYLGLYCEROPHOSPHOCHOLINE ESTERASE) (1- ALKYL-2-ACETYLGLYCEROPHO...			264607, 265008, 265009, 265010, 18108351, 284689, 265021, 18108370, 18108374, 18108385
924	21431341 (1847, 1848)	Novel Protein sim. GBank gi 2497888 sp Q06983 PAFA_MOUSE - PLATELET- ACTIVATING FACTOR ACETYLHYDROLASE PRECURSOR (PAF ACETYLHYDROLASE) (PAF 2- ACYLHYDROLASE) (LDL-ASSOCIATED PHOSPHOLIPASE A2) (LDL-PLA2) (2-ACETYL-1- ALKYLGLYCEROPHOSPHOCHOLINE ESTERASE) (1- ALKYL-2-ACETYLGLYCEROPHO...		esterase	264510
925	20630332 (1849, 1850)	Novel Protein sim. GBank gi 2497888 sp Q06983 PAFA_MOUSE - PLATELET- ACTIVATING FACTOR ACETYLHYDROLASE PRECURSOR (PAF ACETYLHYDROLASE) (PAF 2- ACYLHYDROLASE) (LDL-ASSOCIATED PHOSPHOLIPASE A2) (LDL-PLA2) (2-ACETYL-1- ALKYLGLYCEROPHOSPHOCHOLINE ESTERASE) (1- ALKYL-2-ACETYLGLYCEROPHO...			284603
926	78097657 (1851, 1852)	Novel Protein sim. GBank gi 3882325 db BAA34522.1 - (AB018345) KIAA0802 protein (Homo sapiens)			55811937, 263972, 264638
927	37038201 (1853, 1854)	Novel Protein sim. GBank gi 2497888 sp Q06983 PAFA_MOUSE - PLATELET- ACTIVATING FACTOR ACETYLHYDROLASE PRECURSOR (PAF ACETYLHYDROLASE) (PAF 2- ACYLHYDROLASE) (LDL-ASSOCIATED PHOSPHOLIPASE A2) (LDL-PLA2) (2-ACETYL-1- ALKYLGLYCEROPHOSPHOCHOLINE ESTERASE) (1- ALKYL-2-ACETYLGLYCEROPHO...		UNCLASSIFIED	284769

928	80070610 (1855, 1856)	Novel Protein sim. GBank gi 156148 (M30316) - xanthine dehydrogenase (AA at 2538) [Celliphora vicina]		dehydrogenase	264605
929	20830336 (1857, 1858)	Novel Protein sim. GBank gi 4115938 gb AAD03446.1] - (AF116223) No definition line found [Arabidopsis thaliana]		UNCLASSIFIED	264603
930	5496348 (1859, 1860)	Novel Protein sim. GBank gi 4490608 emb CAB38642.1] - (AJ133495) ribonucleotide reductase major subunit [Staphylococcus aureus]			264259
931	10245731 (1861, 1882)	Novel Protein sim. GBank gi 5459396 emb CAB50754.1] - (AL098839) putative integral membrane transport protein [Streptomyces coelicolor]		reductase	264488
932	80420613 (1863, 1864)	Novel Protein sim. GBank gi 5669523 dbj BAA83045.1] - (AB029016) KIAA1093 protein [Homo sapiens]		transport	264508, 264905, 264906, 264908, 264600, 264802, 264803, 264805, 264766, 264558, 18108387, 264486
933	94326010 (1865, 1866)	Novel Protein sim. GBank gi 119111 sp P12978 EBN2_EBV - EBNA-2 NUCLEAR PROTEIN		UNCLASSIFIED	264508, 264666, 264668, 264693, 27486261, 18108370, 65274791, 284636, 264559, 22279002
934	80039105 (1867, 1868)				264389, 22279002
935	80063162 (1869, 1870)				
936	80026632 (1871, 1872)	Novel Protein sim. GBank gi 645686 (M32103) - ORF-27 [Staphylococcus aureus]		UNCLASSIFIED	264604, 264605, 264693, 18108370, 18108374
937	80250273 (1873, 1874)	Novel Protein sim. GBank gi 13806689 pir CGHU1V - collagen alpha 1(V) chain precursor - human		UNCLASSIFIED	22278985, 22278996, 264602, 264687, 32633988, 18108387
938	80028633 (1875, 1876)	Novel Protein sim. GBank gi 2226243 emb CAA74531.1] - (Y14083) hypothetical protein [Bacillus subtilis]		synthase	263978
939	11071684 (1877, 1878)			UNCLASSIFIED	264802
940	94144252 (1879, 1880)	Novel Protein sim. GBank gi 3560166 emb CAA20678] - (AL031525) ubiquitin carboxyl-terminal hydrolase [Schizosaccharomyces pombe]			264600
941	11398414 (1881, 1882)			ubiquitin	264905, 264908, 284907, 264908, 264909, 264511, 264910, 264592, 33657402, 284598, 264758, 264760, 264683, 264766, 264768, 264789, 33657023, 33657109, 264628, 264629, 264630, 264635, 264636, 264555, 264637, 264556, 264638, 264639, 83373044, 16108385, 264565, 16108391
942	13484122 (1883, 1884)				264593
943	80080258 (1885, 1886)	Novel Protein sim. GBank gi 4033729 (AF036595) - apolipoprotein N-acyltransferase [Pseudomonas aeruginosa]		UNCLASSIFIED	264760
944	80216088 (1887, 1888)	Novel Protein sim. GBank gi 2494784 sp Q50729 GUAA_MYCTU - GMP SYNTHASE [GLUTAMINE-HYDROLYZING] (GLUTAMINE AMIDOTRANSFERASE) (GMP SYNTHETASE)		UNCLASSIFIED	264600, 264687, 264689, 264563
945	60052477 (1889, 1890)	Novel Protein sim. GBank gi 732353 sp P39606 YWCH_BACSU - HYPOTHETICAL 36.6 KD PROTEIN IN QOXD-VPR INTERGENIC REGION		synthase	264511, 264603
946	79248402 (1891, 1892)			UNCLASSIFIED	264906, 264804, 264605, 265020, 16108387
					265017

947	81802699 (1893, 1894)	Novel Protein sim. GBank gi 2896770 emb CAA172471 - (AL021899) hypothetical protein RV2033c [Mycobacterium tuberculosis]	Contains protein domain (PF00459) - inositol monophosphatase family	phosphatase	18108394, 22278996, 264907, 264909, 285008, 285009, 264910, 264758, 264800, 284602, 285018, 264605, 264769, 264689, 264893
948	89185538 (1895, 1898)	Novel Protein sim. GBank gi 2827284 (AF041037) - novel antagonist of GGF signaling [Homo sapiens]		lgf	18108398, 56182575, 22278997, 22278999, 60432049, 29331822, 29331828, 264907, 58182435, 55811388, 285011, 264800, 285017, 265018, 265019, 18108351, 265020, 265021, 285022, 27488265, 283972, 55811578, 264638, 80170394, 264566
949	88081768 (1897, 1898)	Novel Protein sim. GBank gi 4507985 ref NP_003427.1 pZNF1 - zinc finger protein 135 (clone pHZ-17)	Contains protein domain (PF00086) - Zinc finger, C2H2 type	transcription factor	29331825, 21908764, 27488281, 21908766, 52644298, 33857349, 87188518, 59994075, 265020, 285021, 87168559, 52644150, 264637
950	79485872 (1899, 1900)	Novel Protein sim. GBank gi 1079461 ref S43865 - Cytochrome b, type II - potato (fragment)	Contains protein domain (PF00038) - Intermediate filament proteins	struct	264683, 18108361
951	20451411 (1901, 1902)	Novel Protein sim. GBank gi 5420387 emb CAB48879.1 - (A243459) proteophosphoglycan [Leishmania major]		UNCLASSIFIED	284804
952	79566954 (1903, 1904)	Novel Protein sim. GBank gi 5305702 pb AAD41779.1 AF12686 - (AF126867) calpain-like protease [Mus musculus]		calhepsin	264910, 264691
953	10186003 (1905, 1906)	Novel Protein sim. GBank gi 2495842 sp Q47142 YFHS_ECOLI - HYPOTHETICAL 41.8 KD PROTEIN IN CSIE-GLYA INTERGENIC REGION		transport	264510
954	9893328 (1907, 1908)	Novel Protein sim. GBank gi 2360965 (AF016253) - D-amino acid dehydrogenase [Klebsiella aerogenes]		dehydrogenase	264508
955	95313410 (1909, 1910)	Novel Protein sim. GBank gi 5454064 ref NP_006319.1 pSIP1 - SYT interacting protein	Contains protein domain (PF00076) - RNA recognition motif (a.k.a. RRM, RBD, or RNP domain)	dna_rna_bind	56994075, 264509, 264905, 264908, 284907, 264908, 284909, 264510, 264910, 264758, 264759, 285010, 264801, 264780, 18108351, 264762, 264763, 264764, 264766, 264686, 264787, 264887, 264788, 264789, 264689, 284628, 264829, 264630, 264631, 264632, 284634, 284635, 264638, 264637, 264638, 58182323, 264639, 18108388, 264563, 264564
956	80064224 (1911, 1912)	Novel Protein sim. GBank gi 2052129 emb CAB08155 - (Z94752) rifM [Mycobacterium tuberculosis]			264605
957	80036208 (1913, 1914)	Novel Protein sim. GBank gi 1709787 sp Q00451 PRF1_LYCES - 38.4 KD PROLINE-RICH PROTEIN		UNCLASSIFIED	264603, 18108362
958	80036448 (1915, 1916)	Novel Protein sim. GBank gi 2131050 emb CAB09260 - (Z95844) opcA [Mycobacterium tuberculosis]		UNCLASSIFIED	264908, 264910, 264762, 263978, 264637
959	80026647 (1917, 1918)	Novel Protein sim. GBank gi 2129478 ref S51939 - chitinase (EC 3.2.1.14) precursor - beet		UNCLASSIFIED	264602, 264692
960	37815408 (1919, 1920)	Novel Protein sim. GBank gi 2129478 ref S51939 - chitinase (EC 3.2.1.14) precursor - beet		UNCLASSIFIED	284259
981	20567383 (1921, 1922)				283978
982	11399318 (1923, 1924)				264593

963	80590374 (1925, 1928)				UNCLASSIFIED	264510, 264288, 264555, 264558, 264559, 264488
964	79832019 (1927, 1928)	Novel Protein sim. GBank gij4598822dbj BAA76833.1 - (AB023208) KIAA0989 protein [Homo sapiens]			UNCLASSIFIED	264112, 264910, 264689
965	91229485 (1929, 1930)	Novel Protein sim. GBank gij5420387 emb CAB46678.1 - (AJ243459) proteophosphoglycan [Leishmania major]			UNCLASSIFIED	264488, 265017, 264448, 264634, 264559, 83373044
966	95028215 (1931, 1932)				UNCLASSIFIED	264908, 264592, 264598, 264604, 264768, 21908784, 264892, 264893, 264829, 264638, 264638
967	79255708 (1933, 1934)	Novel Protein sim. GBank gij1731207 sp Q11156 RGX3_MYCTU - SENSORY TRANSDUCTION PROTEIN REGX3	Contains protein domain (PF00072) - Response regulator receiver domain		phosphatase	264780
968	79560288 (1935, 1938)	Novel Protein sim. GBank gij2661836 emb CAA75187.1 - (Y14964) putative transport protein [Methylophilus methylotrophus]			transport	264693
969	79919470 (1937, 1938)	Novel Protein sim. GBank gij5419878 emb CAB46422.1 - (AL098747) hypothetical protein [Homo sapiens]	Contains protein domain (PF00098) - Zinc finger, C2H2 type		dna_rna_bind	35598286, 264685, 264686, 35695917, 264692, 18108374, 264835
970	95085947 (1939, 1940)				UNCLASSIFIED	18108392, 18108394, 18108398, 22278995, 22278998, 22278998, 22278998, 26147820, 264828, 265008, 265007, 265008, 265009, 18108348, 33109954, 265010, 265011, 18108351, 264288, 21908767, 21906768, 18108370, 18108374, 18108377, 264830, 264835, 18108380, 83373044, 18108387, 18108388
971	78919770 (1941, 1942)				UNCLASSIFIED	265007, 265020, 22279002
972	20710704 (1943, 1944)					264557
973	20370183 (1945, 1946)	Novel Protein sim. GBank gij1723118 sp P53980 Y174_HUMAN - HYPOTHETICAL PROTEIN KIAA0174				264804
974	80057103 (1947, 1948)				UNCLASSIFIED	264565
975	10196018 (1949, 1950)				UNCLASSIFIED	264510
976	80205742 (1951, 1952)	Novel Protein sim. GBank gij3881459 emb CAA92988.1 - (Z68753) predicted using GeneFinder; Similarity to Yeast hypothetical protein YIK9 (SW:YIK9_YEAST); cDNA EST EMBL:D27880 comes from this gene; cDNA EST EMBL:D27879 comes from this gene; cDNA EST EMBL:D84477 comes from this gene...			UNCLASSIFIED	264508, 264908, 264758, 264832, 264639, 264583
977	10355348 (1953, 1954)	Novel Protein sim. GBank gij549458 sp Q05353 XYS3_PSEPU - XYLDLEGF OPERON TRANSCRIPTIONAL ACTIVATOR 3			UNCLASSIFIED	264808
978	80025927 (1955, 1956)				UNCLASSIFIED	264600, 264602, 264603, 264604
979	80447820 (1957, 1958)	Novel Protein sim. GBank gij3171904 emb CAA75888.1 - (Y15906) DIA-12C protein [Homo sapiens]			UNCLASSIFIED	264787, 264788, 265008, 265007, 264908
980	80025928 (1959, 1960)				UNCLASSIFIED	264600, 264602, 264605
981	80098550 (1981, 1982)	Novel Protein sim. GBank gij3589940 (AF017368) - fecaligental dysplasia protein 2 [Mus musculus]			UNCLASSIFIED	264892, 264555, 264556, 264557, 264559

982	80185670 (1963, 1984)	Novel Protein sim. GBank gl 29502020[emb CAA71575] - (Y10545) fused-cdb [Escherichia coli]		UNCLASSIFIED	264404
983	90985041 (1965, 1966)	Novel Protein sim. GBank gl 476389[pr j43402 - myosin heavy chain-B, neuronal - chicken		struct	65274572, 58182575, 264908, 284908, 285007, 265008, 264756, 265010, 55811150, 33657023, 284634, 284557, 284556
984	20466976 (1967, 1968)	Novel Protein sim. GBank gl 3451504[emb CAA07660.1] - (AJ007747) hypothetical protein BbLPS1.21 [Bordetella bronchiseptica]	Contains protein domain (PF00534) - Glycosyl transferases group 1	UNCLASSIFIED	264605
985	85461368 (1969, 1970)	Novel Protein sim. GBank gl 3451504[emb CAA07660.1] - (AJ007747) hypothetical protein BbLPS1.21 [Bordetella bronchiseptica]		transferase	56182435, 284600
986	87102886 (1971, 1972)			UNCLASSIFIED	264106, 284110, 265020, 50170615
987	79867231 (1973, 1974)				284908
988	19858861 (1975, 1976)			UNCLASSIFIED	264800
989	88095329 (1977, 1978)			UNCLASSIFIED	264506, 265017, 264534, 264564
990	88057746 (1978, 1980)	Novel Protein sim. GBank gl 5725508[gb AAD40800.1]AF080152 - (AF080152) METH1 protein [Homo sapiens]	Contains protein domain (PF01421) - Reprolysin (M12B) family zinc metalloprotease	oxidase	284259, 284908, 265008, 284910, 284586, 264389, 284288, 264786, 284626, 284635, 264586
991	10106140 (1981, 1982)	Novel Protein sim. GBank gl 2105040[emb CAB08835] - (Z95436) hypothetical protein Rv3645 [Mycobacterium tuberculosis]	Contains protein domain (PF00211) - Adenylate and Guanylate cyclase catalytic domain	UNCLASSIFIED	264909
992	79845694 (1983, 1984)			UNCLASSIFIED	264506, 284593
993	10814053 (1985, 1986)	Novel Protein sim. GBank gl 3329297 (AE001355) - Ribonucleoside Reductase, Large Chain [Chlamydia trachomatis]		reductase	264807
994	11090590 (1987, 1988)				264602
995	84321911 (1989, 1990)	Novel Protein sim. GBank gl 5106572[gb AAD39760.1]AF14394 - (AF14394) transcriptional activator SRCAP [Homo sapiens]	Contains protein domain (PF00176) - SNF2 and others N-terminal domain	helicase	18108398, 85274572, 22276996, 264490, 60432049, 29331827, 28146496, 284506, 264905, 284907, 284906, 56182435, 265008, 264591, 264592, 60432228, 60431735, 33657402, 264595, 264758, 21906754, 265010, 265017, 265018, 264605, 264780, 264448, 284763, 284766, 21906765, 21906768, 21906789, 55811957, 264692, 264693, 264628, 35696423, 55811576, 35695855, 284838, 264555, 284558, 264558, 83373044, 22278002, 264563
996	91013745 (1991, 1992)	Novel Protein sim. GBank gl 2811719 (AC004227) - KIA001LB [Homo sapiens]	Contains protein domain (PF00595) - PDZ domain (Also known as DHR or GLGF)	phosphatase	83373044, 29331824, 283978, 55611957, 56526486, 87168516, 284810, 284906, 264565, 264586, 264693, 264768
997	80503347 (1993, 1994)	Novel Protein sim. GBank gl 2648101 (AE001001) - ABC transporter, ATP-binding protein [Archaeoglobus fulgidus]	Contains protein domain (PF00005) - ABC transporter	transport	35696286, 22276997, 22278998, 264508, 264905, 264908, 265010, 284600, 284602, 284605, 264686, 284769, 265021, 264585, 16108391
998	11397390 (1995, 1996)	Novel Protein sim. GBank gl 123530[sp P04828]HRPX_PLALO - HISTIDINE-RICH GLYCOPROTEIN PRECURSOR			264595

999	11766047 (1997, 1998)	Novel Protein sim. GBank gi 2508687 sp P46490 YFCA_HAEIN - HYPOTHETICAL PROTEIN H0198		UNCLASSIFIED	264682
1000	20727844 (1999, 2000)			UNCLASSIFIED	264602
1001	86873131 (2001, 2002)	Novel Protein sim. GBank gi 2224689 db BAA20833 - (AB002377) KIAA0379 [Homo sapiens]	Contains protein domain (PF00023) - Ank repeat	kinase	50432049, 264907, 264908, 264511, 264603, 264683, 264684, 264687, 264689, 29148627, 21906769, 264692, 18108385, 22278000 265009, 264369, 265020
1002	80189603 (2003, 2004)	Novel Protein sim. GBank gi 566121 sp P37709 TRHY_RABIT - TRICHOHYALIN		struct	
1003	17833481 (2005, 2006)				265019
1004	16314687 (2007, 2008)	Novel Protein sim. GBank gi 854065 emb CAA58337 - (X63413) U88 [Human herpesvirus 6]			264635
1005	78617144 (2008, 2010)	Novel Protein sim. GBank gi 114073 sp P07672 APT_ECOLI - ADENINE PHOSPHORIBOSYLTRANSFERASE (APRT)		UNCLASSIFIED	264506
1006	37815429 (2011, 2012)			UNCLASSIFIED	264259
1007	78820871 (2013, 2014)	Novel Protein sim. GBank gi 4062979 db BAA36210.1 - (AB017138) epsilon subunit of malonate decarboxylase [Pseudomonas putida]		synthase	264805
1008	86034444 (2015, 2016)	Novel Protein sim. GBank gi 2808807 emb CAA04607.1 - (AJ001206) putative trehalose synthase [Streptomyces coelicolor]		synthase	265007, 264602, 264605, 264750, 264636
1009	57451289 (2017, 2018)	Novel Protein sim. GBank gi 3839077 AF090113 - AMPA receptor binding protein [Rattus norvegicus]	Contains protein domain (PF00595) - PDZ domain (Also known as DHR or GLGF).	kinase	264102, 264286
1010	84672537 (2019, 2020)	Novel Protein sim. GBank gi 3746332 AF016307 - possible NADH-dependent oxidase, may function as a demethylase [Sinorhizobium meliloti]		dehydrogenase	264592
1011	85546916 (2021, 2022)	Novel Protein sim. GBank gi 2342647 gb AA865591.1 - (U90653) DHH-C-domain-containing cysteine-rich protein [Homo sapiens]		UNCLASSIFIED	35686052, 264905, 264905, 264784, 264788, 35685817, 264629
1012	93294458 (2023, 2024)	Novel Protein sim. GBank gi 3413411 emb CAA20272 - (AL031231) guanosine pentaphosphate synthetase/ polyribonucleotide nucleotidyltransferase [Streptomyces coelicolor]	Contains protein domain (PF00013) - KH domain	phosphorylase	35686052, 264905, 264600, 264601, 264602, 264605, 264782, 264788, 264788, 264689
1013	86095772 (2025, 2026)			UNCLASSIFIED	264581, 21908788
1014	86608828 (2027, 2028)				29331824, 265019, 265020

1015	95418879 (2028, 2030)	Novel Protein sim. GBank gij4159995 (AF083095) - SELUL [Mus musculus]	Contains protein domain (PF00040) - Fibronectin type II domain	struct	22278994, 22278995, 56994075, 22278996, 22278999, 264259, 29331825, 29331828, 264907, 56162435, 264510, 264591, 264593, 60433356, 264594, 55812038, 264758, 21908754, 33657084, 265010, 264600, 265017, 265018, 265019, 18106351, 21906765, 21906766, 21906767, 21906768, 55811957, 265022, 33657023, 85274820, 33657182, 32833986, 18108370, 18108377, 55811578, 35696423, 264630, 22279000, 264585
1016	79559694 (2031, 2032)	Novel Protein sim. GBank gij25069694p41407/ACPD, ECOLI - ACYL CARRIER PROTEIN PHOSPHODIESTERASE (ACP PHOSPHODIESTERASE)		esterase	284688, 264693
1017	11069213 (2033, 2034)	Novel Protein sim. GBank gij5103943jdbj/BA479259.1] - (AP000059) 602aa long hypothetical oligopeptide-binding protein oppA [Aeropyrum pernix]	Contains protein domain (PF00496) - Bacterial extracellular solute-binding proteins, family 5	transport	264600
1018	80072430 (2035, 2036)	Novel Protein sim. GBank gij4493973jemb/CAB39032.1] - (AL034559) predicted using hexExon; MAL3P7.14 (PF00925w), Hypothetical protaln, len: 488 aa [Plasmodium falciparum]			22278996, 26148627, 264583
1019	11703807 (2037, 2038)			UNCLASSIFIED	264886
1020	80234432 (2039, 2040)				264508, 264509, 264512, 264800, 264782, 264769, 264889, 18108370, 264638, 264638, 264486
1021	37036243 (2041, 2042)	Novel Protein sim. GBank gij4633407jbb/AAD28659.1/AF127795 (trehalase biosynthetic enzyme TreY [Rhizobium leguminosarum bv. viciae])		synthase	264789
1022	80502627 (2043, 2044)	Novel Protein sim. GBank gij1761230jemb/CAB08271] - (Z63867) hypothetical protein Rv3137 [Mycobacterium tuberculosis]	Contains protein domain (PF00459) - Inositol monophosphatase family	phosphatase	35696052, 264508, 265008, 265009, 264769, 18108387, 264583
1023	11399341 (2045, 2046)	Novel Protein sim. GBank gij3777485 (U82083) - calcium transporting ATPase [Pichia angustia]	Contains protein domain (PF00122) - E1-E2 ATPase	ATPase, associated	264593
1024	80037129 (2047, 2048)			UNCLASSIFIED	52648842, 33657402, 33657023, 18108379, 55811578, 264631, 264558, 264557, 264559, 18108388, 264566
1025	79844200 (2049, 2050)	Novel Protein sim. GBank gij348304jemb/CAA20556] - (AL031371) putative transport system permease protein [Streptomyces coelicolor]		transport	264693
1026	80025946 (2051, 2052)	Novel Protein sim. GBank gij1174922jpp/Q2322JUVRD, HAEIN - DNA HELICASE II		helicase	264802
1027	17658234 (2053, 2054)	Novel Protein sim. GBank gij475772jpe/NP_004866.1/pAGTA - angiotensin/vasopressin receptor ALLAVP-like		UNCLASSIFIED	285017

1028	20297828 (2055, 2056)	Novel Protein sim. GBank gij2791409[emb]CAA16003] - (AL021184) acn [Mycobacterium tuberculosis]	Contains protein domain (PF00330) - Aconitase family (aconitate hydratase)	UNCLASSIFIED	284600
1029	94655080 (2057, 2058)			UNCLASSIFIED	284595
1030	88095343 (2059, 2060)			UNCLASSIFIED	284907, 284906, 284510, 284512, 285009, 285010, 285011, 284800, 284602, 284803, 284605, 284769, 18106372, 18108374
1031	95289117 (2061, 2062)			UNCLASSIFIED	284905, 284908, 284909, 284595, 284892, 284830, 284634, 284638
1032	94673275 (2063, 2064)	Novel Protein sim. GBank gij4503885[ref]NP_000145.1pGALK - galactokinase 1		UNCLASSIFIED	284689
1033	86464818 (2065, 2066)	Novel Protein sim. GBank gij2882890 (AE000682) - hypothetical protein [Aquifex aeolicus]		UNCLASSIFIED	35896032, 284908, 284510, 18106354, 284687, 284769, 284889, 60431802, 18108385, 284488
1034	78245837 (2067, 2068)	Novel Protein sim. GBank gij405885 (U00007) - methionyl-RNA synthetase [Escherichia coli]		UNCLASSIFIED	284906
1035	78956355 (2069, 2070)			UNCLASSIFIED	284892
1036	65904898 (2071, 2072)			UNCLASSIFIED	284905, 66712502, 284906, 284768
1037	67996058 (2073, 2074)			UNCLASSIFIED	29331824, 284908, 80433436, 265019
1038	20481015 (2075, 2076)	Novel Protein sim. GBank gij2605987 (AF030027) - 24 [Equine herpesvirus 4]	Contains protein domain (PF01477) - PLATLH2 domain		284804, 284634
1039	67260021 (2077, 2078)	Novel Protein sim. GBank gij2352095 (U97023) - DNA topoisomerase I [Fenvidobacterium islandicum]	Contains protein domain (PF01131) - Prokaryotic DNA topoisomerase	UNCLASSIFIED	284092, 284093, 284094, 284683, 284689, 263967
1040	80025840 (2079, 2080)	Novel Protein sim. GBank gij3258555[db]BAA29218.1] - (AP000001) 301aa long hypothetical 2-phosphoglycerate kinase [Pyrococcus horikoshii]	kinase		284907
1041	10156682 (2081, 2082)	Novel Protein sim. GBank gij2058299[emb]CAA66853] - (X98309) ARI protein [Drosophila melanogaster]			284605
1042	11084375 (2083, 2084)	Novel Protein sim. GBank gij1870167[emb]CAA70125] - (Y08921) msK [Streptomyces reticuli]	Contains protein domain (PF00005) - ABC transporter		284565, 284567
1043	80057136 (2085, 2086)	Novel Protein sim. GBank gij5688890[emb]CAB52053.1] - (AL108732) hypothetical protein [Streptomyces coelicolor A3(2)]		UNCLASSIFIED	265006, 284602, 265017
1044	80025952 (2087, 2088)				29331825, 284637
1045	52415462 (2089, 2090)	Novel Protein sim. GBank gij854065[emb]CAA58337] - (X83413) U88 [human herpesvirus 8]	helicase		284686
1046	11754862 (2091, 2092)	Novel Protein sim. GBank gij4210471[db]BAA74535.1] - (AB019033) orf5A [Pseudomonas sp.]		UNCLASSIFIED	284769
1047	37036258 (2093, 2094)	Novel Protein sim. GBank gij3413419[emb]CAA20279] - (AL031232) hypothetical protein SC10H5.07 [Streptomyces coelicolor]	Contains protein domain (PF00220) - Neurohypophysial hormones, N-terminal Domain	UNCLASSIFIED	284687
1048	78186400 (2095, 2096)	Novel Protein sim. GBank gij5051838[db]JAD38328.1[AF07372] - (AF07372) EH domain-binding mitotic phosphoprotein [Homo sapiens]		UNCLASSIFIED	284905, 284634
1049	61755108 (2097, 2098)				
1050	78471521 (2099, 2100)			UNCLASSIFIED	284686

1051	80475471 (2101, 2102)				UNCLASSIFIED	18108374, 264769, 265010, 265011, 264601, 265009, 264604, 264605, 264836, 18108351, 264892
1052	82442962 (2103, 2104)	Novel Protein sim. GBank gi 3123275 sp P35138 ISERA_BACSU - D-3- PHOSPHOGLYCERATE DEHYDROGENASE (PGDH)		Contains protein domain (PF00389) - D-isomer specific 2-hydroxyacid dehydrogenases	dehydrogenase	264508, 264762, 264687, 264486
1053	94851640 (2105, 2106)	Novel Protein sim. GBank gi 5441319 emb CAB46717.1 - (AL034398) dJ1158B12.1 (zinc finger, X-linked, duplicated A) [Homo sapiens]				264668, 18108374, 26331824, 83373044, 21906754, 52645158, 58182435, 264889, 26331827, 27486281, 35898052, 21908785, 35898423, 21908788, 58182575, 21908789, 55811957, 87168518, 35696286, 22278997, 285020, 285011, 265021, 265022, 285007, 285018, 22279000, 22279002, 284482, 264906, 52844150, 264909, 284288, 28331822, 52845080, 264768
1054	79580225 (2107, 2108)				UNCLASSIFIED	264688
1055	80394138 (2109, 2110)	Novel Protein sim. GBank gi 5052508 gb AAD38594.1 AF14560 - (AF145609) BcDNA GH02833 [Drosophila melanogaster]		Contains protein domain (PF00270) - DEAD/DEAH box helicase	helicase	264907, 264802, 264681, 264288, 21908788, 33857109, 55810764, 35695855, 264631
1056	17882319 (2111, 2112)	Novel Protein sim. GBank gi 3021676 db BAA35358 - (D86033) RNA polymerase sigma-70 factor [Pseudomonas fluorescens]			mapolymerase	264906
1057	85667218 (2113, 2114)	Novel Protein sim. GBank gi 1226281 U50308 - No definition line found [Caenorhabditis elegans]			UNCLASSIFIED	264682
1058	80376576 (2115, 2118)					264764
1059	94882754 (2117, 2118)	Novel Protein sim. GBank gi 1170018 sp P46808 GREM_MYCLE - TRANSCRIPTION ELONGATION FACTOR GREM (TRANSCRIPT CLEAVAGE FACTOR GREM)			transcriptfactor	35695052, 35695855, 285009, 264636
1060	79481169 (2119, 2120)	Novel Protein sim. GBank gi 2495087 sp Q09332 UGGG_DROME - UDP- GLUCOSE:GLYCOPROTEIN GLUCOSYLTRANSFERASE PRECURSOR (DUGT)			glycoprotein	29146499, 264881, 284883, 264887
1061	11034025 (2121, 2122)	Novel Protein sim. GBank gi 80254 pir J28334 - protein- lysine-phosphatase (EC 3.1.3.48) Ly-5 precursor (B-cell variant) - mouse			phosphatase	264634
1062	39567937 (2123, 2124)	Novel Protein sim. GBank gi 3334200 sp Q49954 GCSP_SOLITU - GLYCINE DEHYDROGENASE (DECARBOXYLATING) PRECURSOR (GLYCINE DECARBOXYLASE) (GLYCINE CLEAVAGE SYSTEM P-PROTEIN)			dehydrogenase	264593
1063	8490481 (2125, 2126)	Novel Protein sim. GBank gi 2499968 sp Q41228 PSE1_NICSY - PHOTOSYSTEM I REACTION CENTRE SUBUNIT IV A PRECURSOR (PSI-E A)				264508
1064	78881783 (2127, 2128)	Novel Protein sim. GBank gi 82654 pir JUA0088 - 10K zein precursor - maize				265007, 265008, 18108351, 18108385

1065	80021208 (2128, 2130)	Novel Protein sim. GBank gi 2120888 pir S70662 - glycoyltransferase homolog - Bordetella pertussis		transferase	264600, 264602, 264689
1066	17896878 (2131, 2132)	Novel Protein sim. GBank gi 2506382 sp P15042 DNLJ_EC01 - DNA LIGASE (POLYDEOXYRIBONUCLEOTIDE SYNTHASE (NAD+))		synthase	265009
1067	10132178 (2133, 2134)				264809
1068	82062057 (2135, 2136)	Novel Protein sim. GBank gi 4007669 emb CAA223351 - (AL034443) putative oxidoreductase [Streptomyces coelicolor]	Contains protein domain (PF00246) - Aldo/keto reductase family	reductase	264688, 18108382, 264558, 264600, 264760
1069	63002054 (2137, 2138)	Novel Protein sim. GBank gi 4589484 dbj BAA76770.1 - (AB023143) KIAA0928 protein [Homo sapiens]		UNCLASSIFIED	264804
1070	82101992 (2139, 2140)	Novel Protein sim. GBank gi 120304 sp P15932 FLOK_SALTY - FLAGELLAR HOOK-ASSOCIATED PROTEIN 1 (HAP1)		UNCLASSIFIED	264804, 264780
1071	20710568 (2141, 2142)	Novel Protein sim. GBank gi 1750397 (U61261) - glutamate synthase large subunit [Pseudomonas aeruginosa]		synthase	264802
1072	82356540 (2143, 2144)			UNCLASSIFIED	264687, 264688, 21908764, 35686052, 35695917, 35695855, 264600, 264601, 264602, 285008, 264605, 264508, 264905, 264680, 264906, 264762, 264626, 264786
1073	79814400 (2145, 2146)				264909
1074	60105982 (2147, 2148)	Novel Protein sim. GBank gi 477532 pir J48175 - Motch B protein - mouse (fragment)	Contains protein domain (PF00006) - EGF-like domain	synthase	264906
1075	81650293 (2149, 2150)	Novel Protein sim. GBank gi 3893109 emb CAA76940 - (Y17820) CALO protein [Drosophila melanogaster]		UNCLASSIFIED	55994075, 22276896, 264594, 264757, 264596, 265018, 265019, 264681, 264389, 264688, 285020, 18108384, 18108374
1076	80477284 (2151, 2152)	Novel Protein sim. GBank gi 1176203 sp P46442 YHGM_EC01 - HYPOTHETICAL 43.1 KD PROTEIN IN RPLM-HHOA INTERGENIC REGION (F375)		ATPase-associated	264769
1077	79831334 (2153, 2154)	Novel Protein sim. GBank gi 4033487 sp Q44472 TUD4_AGRV1 - PUTATIVE HYDROXYPYRUVATE REDUCTASE		kinase	264805
1078	20286874 (2155, 2156)				
1079	80494518 (2157, 2158)	Novel Protein sim. GBank gi 3413828 emb CAA20296 - (AL031260) hypothetical protein SC9A10.09 [Streptomyces coelicolor]		UNCLASSIFIED	264600
1080	11767168 (2159, 2160)			UNCLASSIFIED	16108394, 264769, 264634, 264636
1081	94747060 (2161, 2162)			UNCLASSIFIED	264684
1082	81490656 (2163, 2164)			UNCLASSIFIED	63373044, 265019, 22278002, 264482, 16108351, 264682, 264906, 264683, 264467, 264758, 264766, 264769, 21908767, 264511, 264910, 264634, 264835, 264905, 264636, 264806, 264637, 264807, 264906, 264784, 264836, 20281099, 264766, 264595

1083	87446717 (2165, 2166)	Novel Protein sim. GBank gi1728451sp Q10523 Y01N_MYCTU - HYPOTHETICAL 44.8 KD PROTEIN CY427.23		UNCLASSIFIED	60424179, 264905, 264908, 264510, 60432228, 264758, 87166474, 264605, 264789, 264689, 16108364, 16106378, 35955855, 264638 264789
1084	37798306 (2167, 2168)	Novel Protein sim. GBank gi1418384sp P32057 WCAI_ECOLI - PUTATIVE COLANIC ACID BIOSYNTHESIS GLYCOSYL TRANSFERASE WCAI		UNCLASSIFIED	
1085	66475368 (2169, 2170)	Novel Protein sim. GBank gi1699190 (U90204) - heat shock protein 80 [Tsukamurella tyrosolvensis]	Contains protein domain (PF00118) - TCP-1/cpn60 chaperonin family	eph	60432228, 264687
1086	79608269 (2171, 2172)	Novel Protein sim. GBank gi1172958sp P46178 RL14_ACYKS - 50S RIBOSOMAL PROTEIN L14	Contains protein domain (PF00236) - Ribosomal protein L14	ribosomalprot	264486
1087	76603979 (2173, 2174)	Novel Protein sim. GBank gi14160198 emb CAA15431 - (AL008563) dJ327 J18.3 (novel CHROMObox family protein) [Homo sapiens]	Contains protein domain (PF00385) - 'chromo' (CHROMatin Organization Modifier) domain	helicase	26331827, 264693
1088	76854963 (2175, 2176)	Novel Protein sim. GBank gi12983155 (AE000693) - phosphoglucosyltransferase/phosphomannomutase [Aquifex aerophilus]		UNCLASSIFIED	264905, 264601, 18108367
1089	80216600 (2177, 2178)	Novel Protein sim. GBank gi14981768 gb AAD36290.1 AE00177 - (AE00177) NADH dehydrogenase, 30 kDa subunit, putative [Thermotoga maritima]	Contains protein domain (PF00329) - Respiratory-chain NADH dehydrogenase, 30 kD subunit	UNCLASSIFIED	264488, 264511, 265011, 264682, 264788, 264889, 21908764, 35695917, 265020, 32633986, 18108370, 35695855
1090	11083625 (2179, 2180)	Novel Protein sim. GBank gi14007680 emb CAA22366 - (AL034443) putative oxidoreductase [Streptomyces coelicolor]			264604
1091	12917471 (2181, 2182)	Novel Protein sim. GBank gi12495562 sp P77239 YLCD_ECOLI - HYPOTHETICAL 44.3 KD PROTEIN IN NFRB-PHEP INTERGENIC REGION PRECURSOR		UNCLASSIFIED	264637
1092	80252266 (2183, 2184)	Novel Protein sim. GBank gi12980099 emb CAA17988.1 - (AL022121) nt1 [Mycobacterium tuberculosis]	Contains protein domain (PF00730) - Endonuclease III	nuclease	264568 264768, 35695917, 35695855, 264600, 264602, 264803, 264605, 18108351 264686
1093	80486304 (2185, 2186)	Novel Protein sim. GBank gi11001642 dbj BAA10373 - (D84002) dGTP triphosphohydrolase [Synecocystis sp.]		UNCLASSIFIED	
1094	10680972 (2187, 2188)	Novel Protein sim. GBank gi14565567 emb CAB40855.1 - (AL049828) putative adenine glycosylase [Streptomyces coelicolor]	Contains protein domain (PF00455) - Bacterial regulatory proteins, deoR family	nuclease	264906, 285007, 284595, 284800, 264602, 264603, 264604, 264605, 284762, 264766, 264789, 264638, 284558, 18108387, 60432113, 284482, 264486
1095	87457250 (2189, 2190)	Novel Protein sim. GBank gi115001 sp P18206 BIOB_BACSH - BIOTIN SYNTHASE (BIOTIN SYNTHETASE)		synthase	264600, 264602, 264603, 264604, 264605, 35695917, 264692, 264631
1096	80025977 (2191, 2192)	Novel Protein sim. GBank gi1114135 sp P08205 ARGA_ECOLI - AMINO-ACID ACETYLTRANSFERASE (N-ACETYLGLUTAMATE SYNTHASE) (AGS)		UNCLASSIFIED	265019 264687
1097	79239560 (2193, 2194)	Novel Protein sim. GBank gi1114135 sp P08205 ARGA_ECOLI - AMINO-ACID ACETYLTRANSFERASE (N-ACETYLGLUTAMATE SYNTHASE) (AGS)		synthase	
1098	79186424 (2195, 2196)	Novel Protein sim. GBank gi1114135 sp P08205 ARGA_ECOLI - AMINO-ACID ACETYLTRANSFERASE (N-ACETYLGLUTAMATE SYNTHASE) (AGS)			

1069	38523638 (2197, 2198)	Novel Protein sim. GBank gi 3915144 sp O33017 TRIMD_MYCLE - TRNA (GUANINE- NI)-METHYLTRANSFERASE (M1G- METHYLTRANSFERASE) (TRNA [GM37] METHYLTRANSFERASE)		UNCLASSIFIED	264803
1100	85738571 (2189, 2200)	Novel Protein sim. GBank gi 3023255 sp Q84420 ACOD_MESAU - ACYL-COA DESATURASE (STEAROYL-COA DESATURASE) (FATTY ACID DESATURASE) (DELTA9)-DESATURASE)	desaturase		264258, 264638
1101	80491857 (2201, 2202)	Novel Protein sim. GBank gi 1174735 sp P43012 TOP1_HAEN - DNA TOPOISOMERASE I (OMEGA-PROTEIN) (RELAXING ENZYME) (UNTWISTING ENZYME) (SWIVELASE)	isomerase Contains protein domain (PF01396) Topoisomerase DNA binding C4 zinc finger		264769
1102	79777814 (2203, 2204)	Novel Protein sim. GBank gi 1806596 U81788 - kinesin-73 [Drosophila melanogaster]	UNCLASSIFIED		264910, 264909
1103	81897259 (2205, 2206)	Novel Protein sim. GBank gi 2835448 AF048976 - synaptic ras GTPase-activating protein p135 Syngap [Rattus norvegicus]	struct		264757
1104	95003115 (2207, 2208)	Novel Protein sim. GBank gi 2835448 AF048976 - synaptic ras GTPase-activating protein p135 Syngap [Rattus norvegicus]	UNCLASSIFIED		264558, 22278002
1105	80255121 (2209, 2210)	Novel Protein sim. GBank gi 2835448 AF048976 - synaptic ras GTPase-activating protein p135 Syngap [Rattus norvegicus]	UNCLASSIFIED		264556
1106	79314110 (2211, 2212)	Novel Protein sim. GBank gi 2835448 AF048976 - synaptic ras GTPase-activating protein p135 Syngap [Rattus norvegicus]	UNCLASSIFIED		264555, 264368
1107	80470019 (2213, 2214)	Novel Protein sim. GBank gi 2835448 AF048976 - synaptic ras GTPase-activating protein p135 Syngap [Rattus norvegicus]	UNCLASSIFIED		264808, 264769
1108	80440816 (2215, 2216)	Novel Protein sim. GBank gi 1173421 sp P43416 SECY_STRSC - PREPROTEIN TRANSLOCASE SECY SUBUNIT	transport		264907, 264510, 264511, 264800, 264802, 264605, 264768, 264769
1109	80084615 (2217, 2218)	Novel Protein sim. GBank gi 2995310 emb CAA18338 - (AL022268) putative ATP-dependent helicase [Streptomyces coelicolor]	helicase		264802, 264605, 264638
1110	80503554 (2218, 2220)	Novel Protein sim. GBank gi 2995310 emb CAA18338 - (AL022268) putative ATP-dependent helicase [Streptomyces coelicolor]			264806, 264593, 265010, 264601, 264603, 264804, 264805, 264882, 264769, 264893, 264838
1111	80071744 (2221, 2222)	Novel Protein sim. GBank gi 2622039 AE000668 - type I resistion modification system, subunit S [Methanobacterium thermoautotrophicum]			18108370, 264557
1112	85010088 (2223, 2224)	Novel Protein sim. GBank gi 3216376 emb CAA18628 - (AL023862) putative oxidoreductase [Streptomyces coelicolor]	UNCLASSIFIED		264808
1113	82456352 (2225, 2226)	Novel Protein sim. GBank gi 1083428 pir J554876 - NAD(P)+ transhydrogenase (B-specific) (EC 1.8.1.1) precursor - mouse	dehydrogenase		264600, 264602, 264604, 264805, 264762, 264768, 264555
1114	14986014 (2227, 2228)	Novel Protein sim. GBank gi 1083428 pir J554876 - NAD(P)+ transhydrogenase (B-specific) (EC 1.8.1.1) precursor - mouse			264836
1115	11765563 (2229, 2230)	Novel Protein sim. GBank gi 1083428 pir J554876 - NAD(P)+ transhydrogenase (B-specific) (EC 1.8.1.1) precursor - mouse	UNCLASSIFIED		264688
1116	79841152 (2231, 2232)	Novel Protein sim. GBank gi 1083428 pir J554876 - NAD(P)+ transhydrogenase (B-specific) (EC 1.8.1.1) precursor - mouse	UNCLASSIFIED		264808

1117	95305465 (2233, 2234)	Novel Protein sim. GBank gl 3255965 emb CAA94089 - (Z70200) U5 snRNP-specific 200kD protein [Homo sapiens]	Contains protein domain (PF00270) - helicase DEAD/DEAH box helicase	18108392, 264486, 263994, 264489, 16108398, 56182575, 22278995, 22278996, 35698286, 58994075, 22278997, 22278998, 22278999, 284258, 29331822, 56182181, 29331824, 66714117, 29331825, 29331826, 60432289, 29331827, 29331828, 35698052, 33656970, 28146498, 264508, 264805, 264508, 264908, 264907, 284908, 66712502, 264511, 265008, 285007, 264512, 265008, 265009, 264910, 60170831, 264591, 264592, 264593, 60433356, 264594, 60433438, 264595, 55612036, 264598, 264756, 264759, 21908754, 33109854, 33657084, 265011, 87168559, 264600, 264801, 264602, 264603, 265017, 264604, 285018, 264605, 265019, 264760, 55811150, 264661, 264762, 18108351, 264446, 264882, 264784, 264883, 264288, 264389, 264884, 264885, 264788, 264787, 264886, 264687, 264768, 264769, 264688, 16108359, 264869, 21908765, 21908766, 21908767, 21908768, 21908769, 55811857, 29148784, 35695917, 265020, 285021, 285022, 60170815, 264680, 264691, 33657023, 284692, 264693, 65274820, 33657109, 33657182, 27486261, 27486262, 27486264, 33657349, 27486265, 35695763, 264691
1118	79583326 (2235, 2236)		UNCLASSIFIED	264691
1119	79842463 (2237, 2238)	Novel Protein sim. GBank gl 5420387 emb CA846679.1 - (AJ243459) proteophosphoglycan [Leishmania major]	UNCLASSIFIED	264907
1120	79480463 (2239, 2240)	Novel Protein sim. GBank gl 1644450 (U67664) - MEX-3 [Caenorhabditis elegans]	collagen	29331827, 265018, 265019, 264681, 265021, 60170815, 16108387
1121	79471716 (2241, 2242)		Contains protein domain (PF00013) - KH domain	264683, 264632, 18108388
1122	79456246 (2243, 2244)		UNCLASSIFIED	
1123	79637119 (2245, 2246)	Novel Protein sim. GBank gl 86800 pir J517788 - 3- dehydroquinase synthase (EC 4.6.1.3) - Mycobacterium tuberculosis	UNCLASSIFIED	264639, 264583
1124	79811596 (2247, 2248)		synthase	264893, 27466265
1125	79757061 (2249, 2250)		UNCLASSIFIED	
1126	79758914 (2251, 2252)	Novel Protein sim. GBank gl 138154 sp P03643 VGG_BPHX - MAJOR SPIKE PROTEIN [G PROTEIN] (GPG)	UNCLASSIFIED	264909
1127	11800830 (2253, 2254)		UNCLASSIFIED	264910
1128	6384685 (2255, 2256)	Novel Protein sim. GBank gl 5002704 emb CA844358.1 - (AJ242830) DNA polymerase I [Methylobacterium sp. DM4]	eph	264905, 264809, 264910
			UNCLASSIFIED	264682
			Contains protein domain (PF00478) - polymerase DNA polymerase family A	264511

1129	80422460 (2257, 2258)	Novel Protein sim. GBank gij5689485dbj BAA03028.1) - (AB028997) KIAA1074 protein [Homo sapiens]	Contains protein domain (PF00170) - bZIP transcription factor	UNCLASSIFIED	265011, 264768
1130	79420151 (2259, 2260)	Novel Protein sim. GBank gij4981328gb AAD35881.1 AE00174 - (AE001747) bioY protein [Thermotoga maritima]		UNCLASSIFIED	264595
1131	80055391 (2261, 2262)	Novel Protein sim. GBank gij1841552 (U89336) - unknown [Homo sapiens]		UNCLASSIFIED	35698286, 22278998, 29331828, 264603, 264905, 264559
1132	82062248 (2283, 2284)	Novel Protein sim. GBank gij1841552 (U89336) - unknown [Homo sapiens]		UNCLASSIFIED	22278998, 264906, 265009, 264600, 264602, 264604, 264805, 264780, 32833988, 18108374
1133	17280437 (2285, 2286)			UNCLASSIFIED	265018
1134	80233376 (2287, 2288)			UNCLASSIFIED	264512, 264534
1135	80029393 (2289, 2270)	Novel Protein sim. GBank gij4539171 emb CAB39700.1) - (AL049485) conserved hypothetical protein [Streptomyces coelicolor]		UNCLASSIFIED	264508, 264600, 264602, 264603, 18108376
1136	79842032 (2271, 2272)	Novel Protein sim. GBank gij4982454 gb AAD36931.1 AE001823) ATP-dependent protease LA, putative [Thermotoga maritima]		UNCLASSIFIED	264906, 264908
1137	90931557 (2273, 2274)	Novel Protein sim. GBank gij4972746 gb AAD34768.1) - (AF132160) unknown [Drosophila melanogaster]	Contains protein domain (PF00515) - collagen TPR Domain	UNCLASSIFIED	22278998, 22278999, 35698052, 264907, 265009, 60433358, 264598, 265010, 264448, 264682, 264787, 264889, 265020, 264892, 55811578, 35695855, 264831, 264832, 22279002
1138	79841163 (2275, 2276)	Novel Protein sim. GBank gij731807 sp P38739 YHC8_YEAST - HYPOTHETICAL 63.8 KD PROTEIN IN GUT1-RIM1 INTERGENIC REGION PRECURSOR	strud		264908
1139	79633561 (2277, 2278)	Novel Protein sim. GBank gij3650031 (AC005396) - putative proline-rich cell wall protein [Arabidopsis thaliana]		UNCLASSIFIED	264893
1140	39480358 (2279, 2280)			UNCLASSIFIED	264593
1141	79838019 (2281, 2282)			UNCLASSIFIED	265019, 264893
1142	19635848 (2283, 2284)			UNCLASSIFIED	264631
1143	87762156 (2285, 2286)	Novel Protein sim. GBank gij3928000 emb CAA05880) - (AJ003125) procollagen I N-proteinase [Homo sapiens]	Contains protein domain (PF00090) - oxidase Thrombospondin type 1 domain	UNCLASSIFIED	56182575, 264908, 264600, 264632, 87188518
1144	80086888 (2287, 2288)			UNCLASSIFIED	264635, 264636, 264907, 264593, 264906, 264568, 264909
1145	14610262 (2289, 2290)			UNCLASSIFIED	264112
1146	82062082 (2291, 2292)			UNCLASSIFIED	264769, 264689, 35698286, 264760, 264905, 264488, 264559
1147	80071761 (2293, 2294)	Novel Protein sim. GBank gij2499003 sp P76422 HID_ECOLI - PHOSPHOMETHYL PYRIMIDINE KINASE (HMP-PHOSPHATE KINASE) (HMP-P KINASE)	kinase		264557
1148	80048433 (2295, 2286)	Novel Protein sim. GBank gij2898734 emb CAA17213.1) - (AL021697) hypothetical protein Rv1097c [Mycobacterium tuberculosis]			264591
1149	11607436 (2297, 2298)				264591

1150	61325074 (2288, 2300)	Novel Protein sim. GBank gij2695085 (AF011337) - putative E1-E2 ATPase [Mus musculus]		ATPase-associated	264468, 35698268, 264907, 264908, 264909, 264910, 264593, 264596, 264758, 264784, 264766, 264768, 264693, 264628, 60431850, 264564, 264568, 264567
1151	80070674 (2301, 2302)	Novel Protein sim. GBank gij4324655(gjAA016978) - (AF108191) DNA polymerase III alpha subunit [Streptomyces coelicolor]		polymerase	264595
1152	80235647 (2303, 2304)	Novel Protein sim. GBank gij3874275(embjCAB07311.1) - (Z92825) predicted using GaneFinder. Similarly to Yeast low affinity glucose transporter HXT4 (PS:32467); cDNA EST EMBL:G12555 comes from this gene; cDNA EST yk404c10.3 comes from this gene; cDNA EST yk404c10.5 comes from this...		glycoprotein	264466, 22276996, 264905, 264628, 264486
1153	80027783 (2305, 2306)	Novel Protein sim. GBank gij4240315(djBAA74938.1) - (AB020720) KIAA0913 protein [Homo sapiens]		UNCLASSIFIED	264910, 264555, 264557
1154	83002965 (2307, 2308)	Novel Protein sim. GBank gij586655(jspP37617)ATZN_ECOLI - ZINC-TRANSPORTING ATPASE (ZNU)-TRANSLLOCATING P. TYPE ATPASE	Contains protein domain (PF00122) - E1-E2 ATPase	UNCLASSIFIED	264600, 264636
1155	79411098 (2309, 2310)	Novel Protein sim. GBank gij16480(jspP32139)YIHR_ECOLI - HYPOTHETICAL 34.0 KD PROTEIN IN GLNA-RBN INTERGENIC REGION		UNCLASSIFIED	264688, 264838, 264566
1156	82454917 (2315, 2316)	Novel Protein sim. GBank gij2496461(jspQ50724)Y09S_MYCTU - HYPOTHETICAL 87.3 KD PROTEIN CY78.27C			264908, 264762, 264887, 264769, 264869, 16106374, 35695855
1159	79186451 (2317, 2318)	Novel Protein sim. GBank gij1136406(djBAA11490) - (D79995) similar to pig tubulin-lysine ligase. [Homo sapiens]			264687
1160	91229893 (2319, 2320)	Novel Protein sim. GBank gij2443342(djBAA22380) - (D86764) alpha 2 type I collagen [Rana catesbeiana]			29331822, 29331824, 66714117, 29331826, 60433356, 265018, 265019, 63373044, 16108385, 22279000, 22279002, 264563
1161	7417143 (2321, 2322)	Novel Protein sim. GBank gij4503375(jspNP_001376.1)DPYS - dihydropyrimidinase		UNCLASSIFIED	264368
1162	79635357 (2323, 2324)	Novel Protein sim. GBank gij5052594(jspJAA038807.1)AF145832 - (AF145832) BcDNA.GH06032 [Drosophila melanogaster]		UNCLASSIFIED	264693
1163	79563186 (2325, 2326)	Novel Protein sim. GBank gij5052594(jspJAA038807.1)AF145832 - (AF145832) BcDNA.GH06032 [Drosophila melanogaster]		UNCLASSIFIED	29331827, 264906
1164	79650829 (2327, 2328)	Novel Protein sim. GBank gij5052594(jspJAA038807.1)AF145832 - (AF145832) BcDNA.GH06032 [Drosophila melanogaster]		transport	16108398, 29331827, 29331826, 29146498, 29146499, 16108354, 21908788, 29148827, 21908769, 264683, 16108382, 18108385
1165	80491866 (2329, 2330)	Novel Protein sim. GBank gij4589476(djBAA76766.1) - (AB023138) KIAA0922 protein [Homo sapiens]		UNCLASSIFIED	264602, 264605, 264769, 16108370, 16108374, 264585
1166	88086456 (2331, 2332)	Novel Protein sim. GBank gij4589476(djBAA76766.1) - (AB023138) KIAA0922 protein [Homo sapiens]			264488, 35686286, 22278899, 264258, 88714117, 60432289, 35688052, 264905, 56182435, 265008, 60433436, 264759, 21908764, 33109854, 265017, 265019, 264446, 264288, 264768, 264685, 35696423, 35695855, 264558, 18106385, 60432113

1167	79963882 (2333, 2334)	Novel Protein sim. GBank gl 2560433 dbj BAA23138 - (D76414) ppOpp hydrolase [Staphylococcus aureus]		kinase	264488
1168	88094876 (2335, 2336)			UNCLASSIFIED	264259, 26331827, 56182435, 60433438, 265019, 33657023, 35695855, 264566
1169	11605403 (2337, 2338)			UNCLASSIFIED	264881
1170	21632244 (2338, 2340)			UNCLASSIFIED	264802
1171	20434582 (2341, 2342)	Novel Protein sim. GBank gl 2772914 (AF029249) - precollagen D [Mytilus edulis]		UNCLASSIFIED	264556
1172	79610113 (2343, 2344)	Novel Protein sim. GBank gl 4757646 ref NP_004317.1 pBCL9 - B-cell CLL/lymphoma 9		UNCLASSIFIED	55810784, 35696052, 264634, 264488
1173	80235713 (2345, 2346)	Novel Protein sim. GBank gl 2564053 dbj BAA22946 - (AB007832) Bm trachealless [Bombyx mori]			264508, 264906, 264907, 264909, 264591, 264632, 264636, 264639
1174	20293077 (2347, 2348)	Novel Protein sim. GBank gl 2811027 emb CAA17520 - (AL021958) rmaA [Mycobacterium tuberculosis]		dehydrogenase	264600
1175	20711847 (2349, 2350)	Novel Protein sim. GBank gl 118333 sp P23234 DCIP_ENTCL - INDOLE-3-PYRUVATE DECARBOXYLASE (INDOLEPYRUVATE DECARBOXYLASE)	Contains protein domain (PF00205) - Thiamine pyrophosphate enzymes	carboxylase	264601
1176	80252845 (2351, 2352)	Novel Protein sim. GBank gl 1144520 (U34956) - phosphoribosylformylglycinamide synthase [Mycobacterium tuberculosis]	Contains protein domain (PF00586) - AIR synthase related protein	synthase	264508, 264905, 264593, 264802, 264605
1177	80064847 (2353, 2354)	Novel Protein sim. GBank gl 119791 sp P28643 FABG_CUPLA - 3-OXOACYL-ACYL-CARRIER (PROTEIN) REDUCTASE PRECURSOR (3-KETOACYL-ACYL CARRIER PROTEIN REDUCTASE)	Contains protein domain (PF00106) - short chain dehydrogenase	reductase	264605
1178	84128841 (2355, 2356)	Novel Protein sim. GBank gl 5031697 ref NP_005594.1 pFIG1 - familial intrahepatic cholestasis 1, (progressive, Byler disease and benign recurrent)	Contains protein domain (PF00122) - ATPase associated E1-E2 ATPase	ATPase	55274572, 16108398, 22278998, 22278999, 29331826, 264508, 264906, 264628, 33657402, 33108954, 264768, 21908785, 21908786, 21908788, 55811957, 33657023, 264629, 55811578, 35698423, 264636, 264556, 56182323, 80432113, 22278000, 22279002
1179	80055575 (2357, 2358)	Novel Protein sim. GBank gl 2960090 emb CAA17988.1 - (AL022121) dppA [Mycobacterium tuberculosis]	Contains protein domain (PF00496) - Bacterial extracellular solute-binding proteins, family 5	transport	264803
1180	11794448 (2359, 2360)	Novel Protein sim. GBank gl 2558614 emb CAA04187 - (AJ001493) dehydroquinase dehydratase [Streptomyces coelicolor]	Contains protein domain (PF01220) - Dehydroquinase class II	synthase	264638
1181	17846362 (2361, 2362)			UNCLASSIFIED	265017
1182	81494284 (2363, 2364)	Novel Protein sim. GBank gl 5420387 emb CABA6679.1 - (AJ243459) proteophosphoglycan [Leishmania major]			265007, 265009, 264564, 264908, 264693
1183	78574044 (2365, 2366)	Novel Protein sim. GBank gl 4091677 (AF061331) - alpha galactosidase precursor [Saccharopolyspora erythraea]			264688, 35696423, 264638, 18108365
1184	52559833 (2367, 2368)	Novel Protein sim. GBank gl 2128476 ref J551939 - chitinase (EC 3.2.1.14) precursor - beet		UNCLASSIFIED	264602
1185	78491185 (2369, 2370)			glycoprotein	263967

1186	20224012 (2371, 2372)			UNCLASSIFIED	264559	
1187	79246634 (2373, 2374)			UNCLASSIFIED	29331825, 285017, 18108351	
1188	79831387 (2375, 2376)	Novel Protein sim. GBank gij2986039 (AF054525) - hypothetical protein [Synectococcus PCC7002]		UNCLASSIFIED	264905, 264906	
1189	79609367 (2377, 2378)				264692	
1190	78930569 (2379, 2380)			UNCLASSIFIED	265016	
1191	80310105 (2381, 2382)			UNCLASSIFIED	264600, 264605, 264764, 35695855, 264638, 264486	
1192	13521641 (2383, 2384)		Contains protein domain (PF01581) - FMRFamide related peptide family		264636	
1193	11103364 (2385, 2386)			UNCLASSIFIED	263978	
1194	78863947 (2387, 2388)	Novel Protein sim. GBank gij854065 (emb CAA56337) - (XG3413) U68 [Human herpesvirus 8]		UNCLASSIFIED	285007, 285008	
1195	20445442 (2389, 2390)	Novel Protein sim. GBank gij1780277 (AE000459) - putative oxidoreductase [Escherichia coli]		UNCLASSIFIED	264605	
1196	13000668 (2391, 2392)				264639	
1197	11392317 (2393, 2394)	Novel Protein sim. GBank gij2487380 (sp Q50715 IMDH_MYCTU - INOSINE-5'-MONOPHOSPHATE DEHYDROGENASE (IMP DEHYDROGENASE) (IMPDH) (IMPD)	Contains protein domain (PF00571) - CBS domain		264594	
1198	95290101 (2395, 2396)				284603	
1199	81882011 (2397, 2398)	Novel Protein sim. GBank gij1708525 (sp P54673 P3K1_DICD1 - PHOSPHATIDYLINOSITOL 3-KINASE 1 (PI3-KINASE) (PTDINS-3-KINASE) (PI3K)			264259, 264757, 33109954, 21906786	
1200	9848860 (2399, 2400)			UNCLASSIFIED	264910	
1201	80503751 (2401, 2402)	Novel Protein sim. GBank gij2499877 (sp P70845 BLMH_RAT - BLEOMYCIN HYDROLASE (BLM HYDROLASE) (BMH)	cathepsin		264768, 264769	
1202	80082633 (2403, 2404)	Novel Protein sim. GBank gij606342 (U18997) - ORF_0622; reading frame open far upstream of start; possible frameshift, linking to previous ORF [Escherichia coli]	ribosomal prot		264600, 264558	
1203	82125373 (2405, 2406)			UNCLASSIFIED	264768, 264769, 35695917, 264910, 264760, 264908, 264907, 264829, 264906, 264909, 264786	
1204	80503916 (2407, 2408)	Novel Protein sim. GBank gij2500728 (sp Q59912 SECY_STRGB - PREPROTEIN TRANSLOCASE SECY SUBUNIT			264905, 264769, 264836	
1205	80053961 (2409, 2410)			UNCLASSIFIED	264366	
1206	80241985 (2411, 2412)			UNCLASSIFIED	264558, 264557, 264556	
1207	79841182 (2413, 2414)				29331824, 264909, 285021, 18108370	
1208	87755217 (2415, 2416)	Novel Protein sim. GBank gij2645560 (AF027854) - Bcl-2 related ovarian killer protein [Rattus norvegicus]	Contains protein domain (PF000452) - Apoptosis regulator proteins, Bcl-2 family		29331824, 264909, 285021, 18108370, 264764, 264683, 264769, 264686, 264689	

1208	78185742 (2417, 2418)	Novel Protein sim. GBank gij1175033spjP44398XYLA_HAEIN - XYLOSE ISOMERASE	Contains protein domain (PF00259) - Xylose isomerase	isomerase	284687, 284688
1210	56426884 (2419, 2420)			UNCLASSIFIED	284697, 284693
1211	94665685 (2421, 2422)			transferase	284591, 284592, 284595
1212	79167829 (2423, 2424)	Novel Protein sim. GBank gij421095iprj1S30686 - hypothetical protein o248 - Escherichia coli Novel Protein sim. GBank gij3880625embjCAB07858] - (293785) predicted using GeneFinder; similar to RNA recognition motif. (aka RRM, RBD, or RNP domain): cDNA EST EMBL: T01682 comes from this gene; cDNA EST EMBL: M75823 comes from this gene; cDNA EST EMBL: D27559 comes from this ge...	Contains protein domain (PF01412) - Putative GTP-ase activating protein for Arf		284689, 283967
1213	78859833 (2425, 2426)	Novel Protein sim. GBank gij228282iprj1505375A - vir gene [Bordetella pertussis]		kinase	284909
1214	10144308 (2427, 2428)	Novel Protein sim. GBank gij5728285gbjAAD48398.1(A12616 - (A126162) HERV-H LTR associating protein 2 [Homo sapiens]		UNCLASSIFIED	284906
1215	80050108 (2429, 2430)	Novel Protein sim. GBank gij2326739embjCAB10953] - (298288) recN [Mycobacterium tuberculosis]		UNCLASSIFIED	285009, 284601, 284602, 284603, 33657109
1216	20438324 (2431, 2432)	Novel Protein sim. GBank gij417328iprjP33036MURA_ENTCL - UDP-N- ACETYLGLUCOSAMINE 1- CARBOXYVINYLTRANSFERASE (ENOYLPIRUVATE TRANSFERASE) (UDP-N-ACETYLGLUCOSAMINE ENOLPYRUVYL TRANSFERASE) (EPT)		transferase	284604
1217	85011344 (2433, 2434)				
1218	11083680 (2435, 2436)	Novel Protein sim. GBank gij1805460dbjBAA09022] - (D50453) homologue of succinate semialdehyde dehydrogenase GcbD of E. coli [Bacillus subtilis]		UNCLASSIFIED	284805, 284807, 284808, 284591, 284766, 284691, 284693, 284629, 284630, 284636, 284564
1219	91218252 (2437, 2438)			dehydrogenase	284601
1220	91241524 (2439, 2440)	Novel Protein sim. GBank gij4240315dbjBAA74938.1] - (AB020720) KIAA0913 protein [Homo sapiens]		oncogene	56181686, 28331822, 60432268, 284601, 284692, 284629
1221	83045055 (2441, 2442)	Novel Protein sim. GBank gij2143888iprj152523 - nucleoporin p82 homodog - rat (fragment)			52844507, 284905, 284909, 285008, 285019, 285020, 52644150, 33657023, 284693, 33657182, 35695763, 284634, 22279000, 22279002, 284482
1222	20711865 (2443, 2444)	Novel Protein sim. GBank gij730805iprjP39653SPHR_SYN7 - ALKALINE PHOSPHATASE SYNTHESIS TRANSCRIPTIONAL REGULATORY PROTEIN SPHR	Contains protein domain (PF00486) - Transcriptional regulatory protein, C terminal	phosphatase	284601
1223	11615647 (2445, 2446)				
1224	80432845 (2447, 2448)	Novel Protein sim. GBank gij1172627iprjP46546IPROB_CORGL - GLUTAMATE 5- KINASE (GAMMA-GLUTAMYL KINASE) (GK)	Contains protein domain (PF01472) - PUA domain	kinase	284593 284593, 284600, 284601, 284603, 284605, 284766, 18106378, 284635, 18108367

171

1243	79770458 (2485, 2488)	Novel Protein sim. GBank gij335587jembjCAA19971j - (AL031124) branched-chain amino acid aminotransferase [Streptomyces coelicolor]		UNCLASSIFIED	18108374, 35695917, 35695855, 285009, 264508, 264909
1244	10284821 (2487, 2488)	Novel Protein sim. GBank gij2870848 (AF051845) - Xin [Mus musculus]		UNCLASSIFIED	264691
1245	80437103 (2489, 2490)	Novel Protein sim. GBank gij4596338jdbjBAA76357.1j - (AB016787) cytochrome o ubiquinol oxidase B [Pseudomonas putida]	Contains protein domain (PF00115) - cytochrome C and Quinol oxidase polypeptide I	oxidase	264788
1246	80050321 (2491, 2492)	Novel Protein sim. GBank gij3581849jembjCAA20805j - (AL031541) putative phenylalanine synthetase beta chain [Streptomyces coelicolor]		UNCLASSIFIED	264604, 264638, 264557, 264564
1247	80084831 (2493, 2494)	Novel Protein sim. GBank gij2821684 (AE000842) - adhesion protein [Methanobacterium thermoautotrophicum]			264758, 264605, 264839
1248	88070353 (2495, 2496)	Novel Protein sim. GBank gij1352403jpp09467f16P HUMAN - FRUCTOSE-1,6-BISPHOSPHATASE (D-FRUCTOSE-1,6-BISPHOSPHATE 1-PHOSPHOHYDROLASE) (FBPASE)	Contains protein domain (PF00316) - Fructose-1,6-bisphosphatase	UNCLASSIFIED	18108392, 264259, 29331826, 264106, 264508, 264907, 264828, 285009, 60433356, 264757, 264758, 21908754, 285010, 285011, 285018, 285019, 264780, 18108351, 18108354, 265021, 18108378, 18108377, 264630, 18108385
1249	80056657 (2497, 2498)	Novel Protein sim. GBank gij2791407jembjCAA16001j - (AL021184) hypothetical protein Rv1473 [Mycobacterium tuberculosis]	transport		264908, 265010, 264600, 264603, 264691, 18108376
1250	12694385 (2499, 2500)	Novel Protein sim. GBank gij112785jpp051003MG1_ECOLI - DNA-3-METHYLADENINE GLYCOSIDASE I (3-METHYLADENINE-DNA GLYCOSYLASE I, CONSTITUTIVE) (TAG I)		UNCLASSIFIED	264689
1251	79850448 (2501, 2502)				
1252	79458087 (2503, 2504)			UNCLASSIFIED	264909
1253	80050121 (2505, 2508)	Novel Protein sim. GBank gij5670178jpbjAAD48816.1jAF18131 - (AF181317) NRAMP manganese transport protein MnaA [Salmonella typhimurium]	glycoprotein		264683, 263978 264600, 264603, 18108378
1254	87716767 (2507, 2508)	Novel Protein sim. GBank gij103160jprjS22125 - finger protein unkempt - fruit fly (Drosophila melanogaster)		UNCLASSIFIED	35698286, 264910, 264784, 264888, 21906787, 55811957, 264692, 284556, 264639
1255	79189728 (2509, 2510)				264836
1256	87889508 (2511, 2512)	Novel Protein sim. GBank gij298535jembjCAA04608.1j - (AJ001206) pep2 [Streptomyces coelicolor]		UNCLASSIFIED	60432289, 264600, 264605, 264784, 264687, 264769, 264888, 27488285, 18108374, 18108378
1257	80201435 (2513, 2514)	Novel Protein sim. GBank gij3193306 (AF069300) - contains similarity to Arabidopsis membrane-associated salt inducible-like protein (GB:AL021637) [Arabidopsis thaliana]		UNCLASSIFIED	264094, 265019
1258	20708150 (2515, 2518)				
1259	80188012 (2517, 2518)			UNCLASSIFIED	264602, 263978
1260	80084806 (2519, 2520)			UNCLASSIFIED	264906, 284448, 284908 264834, 264839

1261	67412802 (2521, 2522)	Novel Protein sim. GBank gl 5889511 dbj BAA03039.1 - (AB029010) KIAA1087 protein [Homo sapiens]	Contains protein domain (PF01699) - Sodium/calcium exchanger protein	cadherin	29331824, 264806, 284909, 264788, 264769, 284889, 284893, 284639, 18108384, 284583, 284634
1262	13504589 (2523, 2524)	Novel Protein sim. GBank gl 95100 pir S21334 - hypothetical protein 4 - Agrobacterium tumefaciens		UNCLASSIFIED	
1263	20710997 (2525, 2526)				284602
1264	60083386 (2527, 2528)	Novel Protein sim. GBank gl 3550958 (AF004840) - CDO [Rattus norvegicus]		strut	284834
1265	80253579 (2528, 2530)				284583
1266	79914804 (2531, 2532)			UNCLASSIFIED	
1267	80558918 (2533, 2534)	Novel Protein sim. GBank gl 1085002 pir S55056 - mitochondrial carrier protein DIF-1 homolog - Caenorhabditis elegans	Contains protein domain (PF00153) - Mitochondrial carrier proteins	transport	284768, 284838, 264638, 284567, 284259, 21908754, 284389
1268	88178473 (2535, 2536)	Novel Protein sim. GBank gl 4688445 emb CAB43370.1 - (AL050289) hypothetical protein [Homo sapiens]	Contains protein domain (PF00583) - Acetyltransferase (GNAT) family	UNCLASSIFIED	18108398, 22278995, 56994075, 60424269, 29331827, 264108, 264512, 265007, 265008, 285009, 284595, 33109954, 33857084, 87168559, 284600, 265018, 265019, 264369, 284688, 21908767, 285020, 52844150, 284691, 33857023, 33857348, 18108374, 264556, 18108385, 60432113, 22279002, 284488
1269	79821946 (2537, 2538)	Novel Protein sim. GBank gl 3334791 emb CAA19939 - (AL031107) hypothetical protein SC5A7.10c [Streptomyces coelicolor]		UNCLASSIFIED	284508, 284905, 264908, 284687, 284893
1270	80031420 (2539, 2540)	Novel Protein sim. GBank gl 2851834 sp Q50591 Y0D1_MYCTU - HYPOTHETICAL 50.0 KD PROTEIN CY1A11.01	Contains protein domain (PF01574) - IMP dehydrogenase / GMP reductase N terminus	- dehydrogenase	285010, 284601
1271	79840489 (2541, 2542)			ATPase associated	35696052, 264908
1272	79482878 (2543, 2544)				284888, 284889
1273	80220315 (2545, 2546)	Novel Protein sim. GBank gl 1855685 emb CAB03731 - (Z81368) hypothetical protein Rv2395 [Mycobacterium tuberculosis]		UNCLASSIFIED	284508, 284638
1274	95010802 (2547, 2548)				
1275	20730763 (2549, 2550)	Novel Protein sim. GBank gl 123728 sp P10413 HTPG_ECOLI - HEAT SHOCK PROTEIN HTPG (HIGH TEMPERATURE PROTEIN G) (HEAT SHOCK PROTEIN C52.5)	Contains protein domain (PF00183) - Hsp90 protein	UNCLASSIFIED - eph	264905, 264908, 264909, 264789, 284602
1276	21148844 (2551, 2552)	Novel Protein sim. GBank gl 2128478 pir S51939 - chitinase (EC 3.2.1.14) precursor - beetle		UNCLASSIFIED	264389
1277	20438195 (2553, 2554)				
1278	11088385 (2555, 2556)	Novel Protein sim. GBank gl 1175473 sp P44555 YAAJ_HAEIN - HYPOTHETICAL PROTEIN H10183		UNCLASSIFIED	264556, 284603
1279	21658756 (2557, 2558)	Novel Protein sim. GBank gl 1928513 (U64318) - ATP synthase subunit beta [Mycobacterium thermorescens]		UNCLASSIFIED	
1280	78310859 (2559, 2560)	Novel Protein sim. GBank gl 4938504 emb CAB43882.1 - (AL078465) putative protein [Arabidopsis thaliana]		synthase	284605
				strut	283878

1281	04323966 (2561, 2562)	Novel Protein sim. GBank gll1136501 (U39546) - surface protein MCA-32 [Rattus norvegicus]	Contains protein domain (PF00047) - immunoglobulin domain	UNCLASSIFIED	29331825, 29331826, 264766, 83373044
1282	87537695 (2563, 2564)	Novel Protein sim. GBank gll3328190 (AF074266) - proto-oncogene AF4 [Mus musculus]		UNCLASSIFIED	265008
1283	20466305 (2565, 2566)	Novel Protein sim. GBank gll3281721[emb]CAB070571 - (Z92770) hypothetical protein Rv0153c [Mycobacterium tuberculosis]		UNCLASSIFIED	264605
1284	20636325 (2567, 2568)	Novel Protein sim. GBank gll3928022 (AF057696) - LspB [Haemophilus ducreyi]			264604
1285	80427330 (2569, 2570)	Novel Protein sim. GBank gll417154[sp]P33126[HSB2_ORYSA - HEAT SHOCK PROTEIN 82]	Contains protein domain (PF00183) - eph		264766, 264889, 263967
1286	20465254 (2571, 2572)	Novel Protein sim. GBank gll2078004[emb]CAB084511 - (Z95207) gorA [Mycobacterium tuberculosis]	reductase		264605, 264639
1287	80417530 (2573, 2574)			UNCLASSIFIED	265011, 264602, 264766, 264687, 264769, 264889, 16108370, 264838, 16108385, 264563
1288	85336101 (2575, 2576)	Novel Protein sim. GBank gll5353510[gb]AAD2161.1[AF08891 - (AF088916) emilin precursor [Homo sapiens]	Contains protein domain (PF00386) - collagen C1q domain		35896052, 264107, 264508, 264509, 264905, 264906, 264907, 264908, 264909, 264510, 264511, 265007, 264512, 264910, 265009, 33657402, 264595, 264758, 265011, 265019, 264760, 16108351, 264681, 264784, 264288, 264885, 264766, 264887, 264768, 264769, 265020, 265021, 264534, 264892, 16108370, 264828, 16108374, 35896423, 264555, 264558, 264557, 264556, 16108385, 264564, 264568, 264567, 264486, 16108391
1289	11613847 (2577, 2578)	Novel Protein sim. GBank gll1169985[sp]P48023[GPCR_LYMST - G-PROTEIN COUPLED RECEPTOR GRL101 PRECURSOR]	Contains protein domain (PF00001) - 7 transmembrane receptor (rhodopsin family)	UNCLASSIFIED	264637
1290	19526027 (2579, 2580)	Novel Protein sim. GBank gll2072674[emb]CAB063051 - (Z95120) rhlE [Mycobacterium tuberculosis]	Contains protein domain (PF00271) - Helicases conserved C-terminal domain	im7	264563
1291	80470286 (2581, 2582)	Novel Protein sim. GBank gll1835755 (U86336) - zinc finger protein Png-1 [Mus musculus]	Contains protein domain (PF01530) - Zinc finger, C2HC type	helicase	265007, 265006, 264789
1292	84723316 (2583, 2584)	Novel Protein sim. GBank gll2129173[pir]F84453 - oxaloacetate decarboxylase (EC 4.1.1.3) alpha subunit - Methanococcus jannaschii	transcript factor		264092, 264259, 26331822, 26331824, 264508, 264808, 264809, 264512, 265006, 265009, 264591, 265019, 264389, 264286, 264668, 264788, 264693, 16108374, 264632, 58182323, 264839, 83373044, 22278002, 264482, 264563
1293	80067536 (2585, 2586)	Novel Protein sim. GBank gll2129173[pir]F84453 - oxaloacetate decarboxylase (EC 4.1.1.3) alpha subunit - Methanococcus jannaschii		UNCLASSIFIED	265008, 55812038, 264389, 264556
1294	82125908 (2587, 2588)	Novel Protein sim. GBank gll544177[emb]CAB46803.11 - (AL096811) putative alcohol dehydrogenase (zinc-binding) [Streptomyces coelicolor A3(2)]	blindep		264602, 264605, 264760, 16108351, 264689, 33657023, 264559
1295	11686851 (2589, 2590)		dehydrogenase		264689

1296	11687904 (2591, 2592)	Novel Protein sim. GBank	Contains protein domain (PF01367) - 5'-3' exonuclease	UNCLASSIFIED	264591, 264639
1297	79639300 (2593, 2594)	Novel Protein sim. GBank	Contains protein domain (PF01367) - 5'-3' exonuclease	UNCLASSIFIED	264683
1298	84239508 (2595, 2596)	Novel Protein sim. GBank	Contains protein domain (PF01367) - 5'-3' exonuclease	struct	18108348, 265017
1299	80255378 (2597, 2598)	Novel Protein sim. GBank	Contains protein domain (PF01367) - 5'-3' exonuclease	transcript	264486, 264906, 264908, 22279002, 264566
1300	80064667 (2599, 2600)	Novel Protein sim. GBank	Contains protein domain (PF01367) - 5'-3' exonuclease	UNCLASSIFIED	264605
1301	17839614 (2601, 2602)	Novel Protein sim. GBank	Contains protein domain (PF01367) - 5'-3' exonuclease	UNCLASSIFIED	264906
1302	95418198 (2603, 2604)	Novel Protein sim. GBank	Contains protein domain (PF01367) - 5'-3' exonuclease		85658542, 265020
1303	9684121 (2605, 2606)	Novel Protein sim. GBank	Contains protein domain (PF01367) - 5'-3' exonuclease		264908
1304	79377196 (2607, 2608)	Novel Protein sim. GBank	Contains protein domain (PF01367) - 5'-3' exonuclease	UNCLASSIFIED	264506
1305	19905899 (2609, 2610)	Novel Protein sim. GBank	Contains protein domain (PF01367) - 5'-3' exonuclease		264566
1306	13069230 (2611, 2612)	Novel Protein sim. GBank	Contains protein domain (PF01367) - 5'-3' exonuclease	UNCLASSIFIED	264836
1307	82201029 (2613, 2614)	Novel Protein sim. GBank	Contains protein domain (PF01367) - 5'-3' exonuclease	UNCLASSIFIED	264907, 264592, 264784
1308	21426814 (2615, 2616)	Novel Protein sim. GBank	Contains protein domain (PF01367) - 5'-3' exonuclease	reductase	264555
1309	79233011 (2617, 2618)	Novel Protein sim. GBank	Contains protein domain (PF01367) - 5'-3' exonuclease	transport	264906, 18108354
1310	20466319 (2619, 2620)	Novel Protein sim. GBank	Contains protein domain (PF01367) - 5'-3' exonuclease	UNCLASSIFIED	264805
1311	87613142 (2621, 2622)	Novel Protein sim. GBank	Contains protein domain (PF01367) - 5'-3' exonuclease		35696286, 28331827, 264908, 265008, 264764, 264766, 264886, 21906787, 21908769, 35695917, 264691, 264693, 22278985, 22278998, 22278999, 264905, 264908, 265011, 265017, 265019, 264887, 21908768, 265020, 265021, 33657023, 22279002, 264564
1312	88061720 (2623, 2624)	Novel Protein sim. GBank	Contains protein domain (PF01367) - 5'-3' exonuclease	dna_bind	22278998, 22278999, 264259, 20281099, 29146488, 264508, 264908, 66712502, 80433356, 80433438, 265011, 265017, 264683, 264288, 21908765, 21908787, 29148627, 21908788, 35695917, 265021, 33657023, 33657109, 18108370, 18108377, 35695855, 80432113, 22279000, 264583, 18108390
1313	81225458 (2625, 2626)	Novel Protein sim. GBank	Contains protein domain (PF01367) - 5'-3' exonuclease	ribosomalprot	264683
1314	56828053 (2627, 2628)	Novel Protein sim. GBank	Contains protein domain (PF01367) - 5'-3' exonuclease	interleukinrecept	264881
1315	84357192 (2629, 2630)	Novel Protein sim. GBank	Contains protein domain (PF01367) - 5'-3' exonuclease	finger	

1318	95361808 (2831, 2632)	Novel Protein sim. GBank gij568407[dbj]BAA02987.1] - (A8028958) KIAA1035 protein [Homo sapiens]		kinase	56182575, 56181686, 20281171, 29331622, 29331624, 80424288, 29331825, 35898052, 52844045, 284591, 60432229, 265018, 265019, 55811150, 56161562, 21908765, 21908787, 21908788, 35895917, 80170815, 33857023, 85274820, 33857109, 35895783, 35895855, 18108387, 87188518, 60432113, 22279002, 284584
1317	88055167 (2633, 2634)	Novel Protein sim. GBank gij4838757[gb]AAD30541.1[AF13491] - (AF134916) semaphorin subclass 4 member G [Mus musculus]		UNCLASSIFIED	264093, 264808, 264908, 264369, 264884
1318	95322893 (2835, 2638)	Novel Protein sim. GBank gij4880204[gb]AAD27567.1[AF11417] - (AF114171) hypothetical protein [Sorghum bicolor]		UNCLASSIFIED	18108392, 18108346, 265011, 265017, 18108359, 18108382, 56182323, 18108385, 22279000
1319	94238546 (2837, 2638)				264908, 264909, 265006, 265008, 264592, 265019, 264768, 58181562, 18108388, 264628, 264629, 18108377, 264638
1320	86603587 (2839, 2640)	Novel Protein sim. GBank gij4240163[dbj]BAA74670.1] - (AB020654) KIAA0847 protein [Homo sapiens]		UNCLASSIFIED	35898286, 55812038, 285018, 21806768, 285020, 283978, 22279002
1321	86676351 (2841, 2642)	Novel Protein sim. GBank gij4888505[emb]CAB43377.1] - (AL050276) hypothetical protein [Homo sapiens]	Contains protein domain (PF00651) - BTB/POZ domain	transcript factor	60432049, 29331828, 264907, 264908, 264909, 264910, 55812038, 264801, 264782, 264764, 264766, 264768, 264769, 264828, 18108374, 264834, 264835, 18108385
1322	87755272 (2843, 2644)	Novel Protein sim. GBank gij5262591[emb]CAB45736.1] - (AL080143) hypothetical protein [Homo sapiens]	Contains protein domain (PF00096) - Zinc finger, C2H2 type	dna_bind	29331828, 264908, 265020, 33657023, 264893, 264404
1323	94845931 (2845, 2646)	Novel Protein sim. GBank gij545951[dbj]BAA62407.1] - (AB029821) phosphotyrosine N-methyltransferase [Homo sapiens]		synthase	65274572, 56894075, 264259, 29331822, 29331827, 284104, 56182435, 87188474, 18108351, 264288, 21908766, 21906787, 35895917, 285020, 264893, 65274791, 58182323, 18108387
1324	87737814 (2847, 2648)	Novel Protein sim. GBank gij5031717[ref]NP_005704.1[gpBP - goodpasture antigen-binding protein]	Contains protein domain (PF01852) - START domain		22278986, 22278986, 29331828, 264905, 264907, 29331830, 264908, 264510, 265008, 264595, 264759, 21908754, 265018, 264288, 264786, 264769, 21906768, 265022, 18108376, 264831, 264632, 264834, 264636, 264838, 264583, 284584, 284585, 264566
1325	94847471 (2849, 2650)	Novel Protein sim. GBank gij3284501 (U64657) - similar to the DPT/Kunitz family of inhibitors; most similar to tissue factor pathway inhibitor precursor [Caenorhabditis elegans]	Contains protein domain (PF00090) - Thrombospondin type 1 domain	protease	35898286, 264905, 264908, 284907, 264908, 264909, 264910, 284593, 33857402, 264758, 85858542, 264760, 264786, 284789, 264891, 35898423
1326	87316289 (2651, 2852)	Novel Protein sim. GBank gij1397275 (U61947) - C06G3.6 gene product [Caenorhabditis elegans]		UNCLASSIFIED	264259, 68712502, 284682, 264683, 264635

1327	95322697 (2853, 2854)	Novel Protein sim. GBank gi 7268325p P39189 ALU2_HUMAN - IIII ALU SUBFAMILY SB WARNING ENTRY IIII	Contains protein domain (PF00279) - Plant lipid transfer protein family	UNCLASSIFIED	18108398, 22278996, 22278997, 22278998, 284091, 284259, 29331822, 29331824, 29331825, 29331828, 29331827, 29331828, 284105, 284905, 56182435, 264112, 265008, 285008, 21908754, 285010, 285011, 285017, 265018, 284881, 284448, 284784, 284684, 284288, 284885, 264768, 284688, 21908767, 21908769, 29148829, 285020, 284890, 284691, 284892, 284693, 283867, 33857109, 33857182, 27486282, 33857349, 18108370, 18108374, 55810784, 35895855, 284834, 58182323, 83373044, 87168518, 80432113, 22279000, 22279002, 284583, 264567, 284486, 22278997, 29331826, 284595, 18108351, 284766, 22279002, 284482, 284587
1328	87753493 (2855, 2856)			UNCLASSIFIED	22278996, 29331827, 284684, 264692, 33657109
1329	87755276 (2657, 2658)	Novel Protein sim. GBank gi 4678224 gb AA028989.1 AC007135 - (AC007135) unknown protein [Arabidopsis thaliana]		UNCLASSIFIED	264259, 29331825, 284512, 265019, 285021, 284555, 284558, 56526486
1330	87727737 (2659, 2660)	Novel Protein sim. GBank gi 437310 (L23504) - nodulin [Medicago truncatula]		UNCLASSIFIED	284259, 29331826, 29331827, 35696052, 29331828, 60170831, 264448, 284686, 21906765, 55811957, 265020, 33857023, 33857109, 263973, 55811578, 35896423, 35895855, 56182323
1331	87378764 (2661, 2662)	Novel Protein sim. GBank gi 4589566 dbj BA476815.1 - (AB023188) KIAA0871 protein [Homo sapiens]		UNCLASSIFIED	65274572, 22278996, 56994075, 22278999, 80432048, 284258, 29331822, 29331826, 60432289, 29331827, 35696052, 52644045, 58182435, 284510, 21908754, 87168559, 265018, 265019, 284448, 284288, 284369, 284686, 21908765, 21908768, 21908767, 21908768, 265020, 265021, 52844150, 33857023, 56182323, 18108387, 80432113, 22279002
1332	94845937 (2663, 2664)	Novel Protein sim. GBank gi 5458516 dbj BA482407.1 - (AB029821) phosphatidylethanolamine N-methyltransferase [Homo sapiens]	synthase		80432048, 284258, 29331822, 29331826, 60432289, 29331827, 35696052, 52644045, 58182435, 284510, 21908754, 87168559, 265018, 265019, 284448, 284288, 284369, 284686, 21908765, 21908768, 21908767, 21908768, 265020, 265021, 52844150, 33857023, 56182323, 18108387, 80432113, 22279002
1333	88098476 (2665, 2666)	Novel Protein sim. GBank gi 5689527 dbj BA483047.1 - (AB029018) KIAA1095 protein [Homo sapiens]		UNCLASSIFIED	60432289, 66712502, 264591, 60433358, 60433438, 55812038, 285010, 284639, 56526486
1334	87592388 (2667, 2668)	Novel Protein sim. GBank gi 2682536 (AF036685) - Similar to protein-tyrosine phosphatase [Caenorhabditis elegans]	phosphatase		264905
1335	87644798 (2669, 2870)	Novel Protein sim. GBank gi 4240285 dbj BA474921.1 - (AB020705) KIAA0898 protein [Homo sapiens]	Contains protein domain (PF00643) - B-box zinc finger.	UNCLASSIFIED	22278998, 22278999, 29331827, 284509, 284511, 265007, 265008, 265009, 60433438, 21908754, 87168558, 285017, 284288, 21908765, 21908767, 21908768, 21908769, 285020, 33857109, 27486284, 18108374, 284558, 284638, 264557, 60170394, 284559, 18108395, 284563

1336	8767890 (2671, 2672)	Novel Protein sim. GBank gi485445 sp P33485 VNUA_P NUCLEAR ANTIGEN				264508, 264905, 264512, 264784, 264693, 264635, 264637
1337	84312042 (2673, 2674)	Novel Protein sim. GBank gi5869471 dbj BAA83019.1 - (AB026890) KIAA1067 protein [Homo sapiens]	UNCLASSIFIED			56182575, 56994075, 22278998, 22278999, 264092, 264259, 60432289, 29331826, 264908, 264908, 264909, 264112, 265008, 265008, 60433356, 55812038, 33657084, 285011, 265017, 265018, 265019, 264682, 264446, 264883, 264389, 264688, 264689, 21908768, 21908769, 285020, 264691, 27488261, 20281089, 18108379, 55811578, 35695855, 56182323, 60432113, 22279002, 264587
1338	80366114 (2675, 2676)		UNCLASSIFIED			29331822, 265010, 264286, 264889, 18108370, 35695855
1339	80249231 (2677, 2678)	Novel Protein sim. GBank gi1176422 (U43194) - rhophilin [Mus musculus]	UNCLASSIFIED			35696052, 264908, 264688, 264556, 264558
1340	88316311 (2679, 2680)					264905, 264907, 67168559, 264784
1341	88101485 (2681, 2682)					264681, 264885, 264686, 264892
1342	80089017 (2683, 2684)	Novel Protein sim. GBank gi5019564 emb CAB44507.1 - (AL035542) dJ994E9.5 (hs6M1-17 (novel 7 transmembrane receptor (rhodopsin family) (olfactory receptor like) protein)) [Homo sapiens]	Contains protein domain (PF00001) - Im7 7 transmembrane receptor (rhodopsin family)			264629
1343	80082862 (2685, 2686)	Novel Protein sim. GBank gi4557543 ref NP_001384.1 pECM2 - extracellular matrix protein 2	Contains protein domain (PF00560) - struct Leucine Rich Repeat			264910, 264686, 264534
1344	20562559 (2687, 2688)	Novel Protein sim. GBank gi2144101 pir J55210 - lricarboxylate carrier - rat (fragment)				253976
1345	91225546 (2689, 2690)		glycoprotein			264909, 60170394
1346	80255717 (2691, 2692)	Novel Protein sim. GBank gi3861052 emb CAA19523 - (AL023643) predicted using GeneFinder; similar to serine/threonine kinase; cDNA EST yk248a12.3 comes from this gene; cDNA EST yk358c10.5 comes from this gene; cDNA EST EMBL:M89047 comes from this gene; cDNA EST yk248a12.5 comes...	Contains protein domain (PF00069) - kinase Eukaryotic protein kinase domain			22278996, 264907, 264681, 264885, 264689, 285020, 264693, 22279000, 22279002, 264566
1347	80417393 (2693, 2694)	Novel Protein sim. GBank gi4504379 ref NP_003856.1 pHG38 - orphan G protein- coupled receptor HG38	Contains protein domain (PF00560) - glycoprotein Leucine Rich Repeat			264908, 264908, 264909, 265006, 264910, 265011, 265017, 264784, 264788, 264787, 264789, 264631, 264634, 264638, 264587, 264486
1348	87352335 (2695, 2696)	Novel Protein sim. GBank gi3399720 dbj BAA32100 - (AB010989) peptidylarginine deiminase type IV [Rattus norvegicus]	UNCLASSIFIED			264486, 264489, 264506, 264509, 264510, 264511, 264512, 264591, 264592, 264601, 264684, 264885, 264789, 264532, 264534, 264555, 264556, 264557, 264558, 22279002, 264486

1349	91225548 (2697, 2698)	Novel Protein sim. GBank gij2144101pirj155210 - tricarboxylate carrier - rat (fragment)		UNCLASSIFIED	52846842, 35696286, 22278996, 22278996, 22278999, 264259, 29331822, 29331824, 35696052, 264509, 264905, 264906, 264907, 264908, 264909, 264511, 265008, 264512, 264910, 60170831, 264591, 60433438, 264757, 21806754, 265017, 265018, 264605, 264760, 264782, 264288, 264768, 264689, 21806765, 21906766, 21906787, 21906788, 55811957, 35695917, 265020, 264534, 264691, 264692, 33657023, 264693, 33657349, 16106374, 16106378, 35696423, 60170394, 22279000, 22278002, 264563, 264564
1350	67083136 (2699, 2700)			UNCLASSIFIED	52846842, 264259, 29331825, 264908, 264511, 264604, 264288, 21806769, 265020, 33657162, 33657349, 16106374, 35695955, 264555, 264556, 16106385, 22278002, 264486
1351	67361327 (2701, 2702)	Novel Protein sim. GBank gij4887239igb1A032248.1) - (AF084584) BAW protein [Fugu rubripes]		UNCLASSIFIED	264906, 264907, 264638
1352	60076386 (2703, 2704)	Novel Protein sim. GBank gij2144101pirj155210 - tricarboxylate carrier - rat (fragment)		UNCLASSIFIED	264693, 263981
1353	95345417 (2705, 2706)				35696286, 60424259, 264905, 264509, 264906, 264907, 264908, 264909, 264511, 264512, 264910, 264758, 264598, 55811366, 265011, 264605, 55811150, 264762, 264764, 264766, 52844229, 56181562, 35695917, 265022, 33657023, 264693, 35695763, 60431528, 264629, 263978, 35696423, 35695955, 264630, 264634, 264635, 264636, 264637, 264638, 264639, 16106385, 264563, 264564
1354	95350845 (2707, 2708)	Novel Protein sim. GBank gij4889108igb1A027763.1(AF077030) [hypothesized 43.2 kDa protein [Homo sapiens]]		UNCLASSIFIED	22278995, 22278999, 26331826, 264906, 265008, 33657402, 21806754, 265011, 87168559, 264664, 264369, 264769, 264689, 21906765, 21906766, 52844150, 33657023, 264692, 264693, 16106374, 83373044, 87168516, 22278000
1355	68260186 (2709, 2710)	Novel Protein sim. GBank gij1469199dbj[BAA09487] - (D50928) The KIAA0138 gene product is novel [Homo sapiens]	Contains protein domain (PF00076) - RNA recognition motif. (a.k.a. RRM, RBD, or RNP domain)	dna_ma_bind	22278996, 22278997, 264259, 66714117, 264511, 21906754, 265010, 264768, 264689, 21906765, 21906766, 21906769, 264532, 27468262, 264629, 264636, 264556, 264638, 264639, 264482, 264484

1358	85513981 (2711, 2712)	Novel Protein sim. GBank gij113865 (U40342) - ninein [Mus musculus]		stud	18108387, 22278895, 22278896, 22278898, 284094, 29331828, 264905, 265006, 265007, 265008, 265010, 265017, 265018, 265019, 264784, 18108354, 264889, 21908785, 265022, 18108364, 35698423, 83373044, 18108367
1357	88260268 (2713, 2714)	Novel Protein sim. GBank gij897683[emb]CAA80330] - (Z50028) phosphatidylcholine transfer protein [Bos taurus]	Contains protein domain (PF01852) - START domain		264259, 26331822, 29331825, 264510, 87188559, 265018, 264448, 264288, 21908785, 21908768, 21908788, 285021, 264883, 18108378
1358	38719455 (2715, 2716)	Novel Protein sim. GBank gij556218 (L36831) - transcription regulator [Mus musculus]			264757
1359	87771843 (2717, 2718)			UNCLASSIFIED	264807, 264808, 264510, 264511, 264512, 18108351, 264764, 264534, 33857023, 18108374, 264634, 264635, 264638, 264639, 18108385, 264488, 264587
1360	87738272 (2719, 2720)	Novel Protein sim. GBank gij2598282[emb]CAA75612] - (Y15417) acetalde-CoA ligase [Coprinus chineus]		synthase	80432289, 264805
1381	87583527 (2721, 2722)	Novel Protein sim. GBank gij568943[db]BAA83005.1] - (AB028976) KIAA1053 protein [Homo sapiens]	Contains protein domain (PF00538) - SAM domain (Sterile alpha motif)	UNCLASSIFIED	35698286, 22278897, 22278899, 284258, 29331828, 264508, 264509, 264905, 264907, 264908, 265007, 285009, 33109954, 21908754, 87188474, 265011, 264781, 264883, 284288, 264786, 264789, 264689, 21908788, 265020, 285021, 33857023, 55811578, 35698423, 264634, 80432113, 22279002, 284482, 284488
1362	85287881 (2723, 2724)	Novel Protein sim. GBank gij568941[db]BAA82988.1] - (AB028960) KIAA1037 protein [Homo sapiens]	Contains protein domain (PF00400) - WD domain, G-beta repeat	eph	58182575, 58181886, 80432049, 264258, 29331822, 58182181, 29331827, 35698052, 29331828, 264905, 264908, 264908, 264595, 55812038, 85858542, 55811550, 264681, 264288, 264388, 58181562, 80431528, 55810784, 35698423, 80431850, 284558
1383	85758476 (2725, 2726)	Novel Protein sim. GBank gij130494 (U35776) - ADP-ribosyltransferase factor 1-directed GTPase activating protein [Rattus norvegicus]	Contains protein domain (PF01412) - Putative GTP-ase activating protein for Arf	UNCLASSIFIED	284488, 29331828, 264807, 264887, 264689, 264693
1364	86179488 (2727, 2728)				60432289, 80433356, 60433438, 87188559, 284603, 18108351, 21908786, 35698423, 80432113
1365	83003108 (2729, 2730)	Novel Protein sim. GBank gij4568562[db]BAA76803.1] - (AB023178) KIAA0859 protein [Homo sapiens]		oncogene	264786
1366	87003282 (2731, 2732)	Novel Protein sim. GBank gij108494[db]SS4495 - hypothetical protein YPR021c - yeast [Saccharomyces cerevisiae]	Contains protein domain (PF00153) - Mitochondrial carrier proteins	transport	285007

1367	67721210 (2733, 2734)	Novel Protein sim. GBank gj4884068[emb]CAB43240.11 - (AL050019) hypothetical protein [Homo sapiens]	Contains protein domain (PF01342) - UNCLASSIFIED SAND domain	264486, 52846642, 52846365, 22278995, 56984075, 35886286, 22278996, 22278999, 22279999, 264259, 29331824, 66714117, 29331825, 60432289, 35896052, 284905, 284907, 264909, 264909, 264510, 265009, 265007, 264512, 264910, 265009, 60170931, 33657402, 55612036, 21908754, 285011, 97169559, 265017, 295019, 19108351, 284446, 284992, 264693, 264289, 284369, 284986, 284767, 264699, 21908765, 21908766, 21908769, 21908769, 55911957, 285020, 265021, 265022, 264534, 60170615, 264690, 264691, 19108362, 33657023, 33657109, 33957349, 284628, 19108370, 19108374, 19108376, 55611576, 35698423, 35698955, 284935, 284555, 284637, 284556, 52844332, 90170394, 294559, 19108381, 19108385, 56528486, 22279000, 284563, 284567
1369	94320076 (2735, 2736)	Novel Protein sim. GBank gj464561[sp]P35269[RB15_RAT] RAS-RELATED PROTEIN RAB-15	Contains protein domain (PF00071) - Ras family oncogene	264259, 29331822, 29331826, 60432289, 29331827, 35698052, 264506, 284905, 264906, 264908, 264909, 264510, 295007, 264910, 60433436, 264759, 65659542, 87168559, 284900, 284901, 294760, 264764, 264765, 264769, 52844229, 294689, 35695917, 265020, 265021, 264631, 264932, 264634, 264937, 52844332, 264559, 264939, 63373044, 264563, 264596, 264489, 264567, 265008, 60432229, 90433356, 33657084, 21906784, 21906769, 264555, 284839, 264559, 264597
1369	88634033 (2737, 2738)	Novel Protein sim. GBank gj2062702 (U90550) - butyrophilin [Homo sapiens]	UNCLASSIFIED	22278999, 264259, 29331822, 29331824, 29331829, 29331827, 35696052, 264508, 284509, 264905, 264906, 264907, 264909, 264909, 264510, 264511, 264512, 264759, 295011, 265019, 264764, 264766, 264769, 21906767, 35695917, 19106362, 35698423, 284632, 264635, 264636, 284555, 264638, 264559, 264639, 16109395, 65274727, 264404, 284593, 284566, 264486
1370	95316910 (2739, 2740)	Novel Protein sim. GBank gj5031623[ref]NP_005823.1[ptKCNM - potassium large conductance calcium-activated channel, subfamily M, beta member 2]	potassium_channel	22278999, 264259, 29331822, 29331824, 29331829, 29331827, 35696052, 264508, 284509, 264905, 264906, 264907, 264909, 264909, 264510, 264511, 264512, 264759, 295011, 265019, 264764, 264766, 264769, 21906767, 35695917, 19106362, 35698423, 284632, 264635, 264636, 284555, 264638, 264559, 264639, 16109395, 65274727, 264404, 284593, 284566, 264486
1371	95336512 (2741, 2742)	Novel Protein sim. GBank gj5032203[ref]NP_005714.1[ptTSPA - tetraspan 5]	Contains protein domain (PF00335) - 4 transmembrane segments integral membrane proteins	22278999, 264259, 29331822, 29331824, 29331829, 29331827, 35696052, 264508, 284509, 264905, 264906, 264907, 264909, 264909, 264510, 264511, 264512, 264759, 295011, 265019, 264764, 264766, 264769, 21906767, 35695917, 19106362, 35698423, 284632, 264635, 264636, 284555, 264638, 264559, 264639, 16109395, 65274727, 264404, 284593, 284566, 264486

1372	80248517 (2743, 2744)	Novel Protein sim. GBank gij1840708j[BAA09334] - (D50685) trans-sialidase [Trypanosoma cruzi]		collagen	263978
1373	80499421 (2745, 2746)			UNCLASSIFIED	264768, 21908765, 21908767, 22276989, 264691, 264910, 55812038, 265010, 264681, 264664
1374	95087036 (2747, 2748)	Novel Protein sim. GBank gij111878j[JC1241 - beta-interferon-induced protein - rat		interferon	264907, 264510, 265007, 264512, 265008, 60432229, 264689, 65274781, 264555, 264558, 264557, 83373044, 60432113
1375	84236842 (2749, 2750)	Novel Protein sim. GBank gij5649176j[AD03500.2] - (AF051155) G beta-like protein GBL [Rattus norvegicus]	Contains protein domain (PF00400) - WD domain, G-beta repeat	kinase	52844507, 52845156, 52846842, 52846365, 56182575, 56181866, 22276986, 58984075, 35896286, 22276987, 22276988, 22276989, 264259, 26331822, 52645060, 26331824, 26331825, 80432289, 29331826, 29331827, 29331828, 35896052, 33856970, 264905, 264508, 264908, 264907, 264908, 29331830, 284908, 265008, 264511, 265007, 265008, 265009, 33857402, 60433358, 52846317, 33108954, 33657084, 52844286, 87168474, 87168559, 264600, 265017, 265018, 265019, 55811150, 18108351, 264448, 264764, 284288, 264369, 264766, 52644229, 21906785, 21906786, 21906787, 21906788, 21908769, 55811957, 35895917, 265020, 265021, 52844150, 33857023, 264683, 85274620, 52845129, 33857109, 27488281, 33857349, 27486285, 35895783, 18108376, 55810764, 35898423, 35895855, 264630, 264631, 264634, 264636, 264555, 264638, 18108385, 87168518, 60432113, 22279000, 22279002, 264583, 284564, 264566, 264567
1376	87398050 (2751, 2752)	Novel Protein sim. GBank gij138350j[P28986]VGLX_HSVB - GLYCOPROTEIN X PRECURSOR		UNCLASSIFIED	264768, 264769, 35895917, 22276987, 264691, 264259, 29331822, 264893, 35896052, 264508, 264509, 264905, 264906, 264626, 264908, 264629, 18108372, 264909, 264510, 264511, 264512, 265008, 264630, 264631, 264910, 264832, 264634, 264635, 264636, 264591, 264592, 264637, 264638, 264558, 264639, 33857402, 264595, 18108365, 56528486, 265010, 265011, 264600, 284583, 264782, 264564, 264565, 264784, 264486, 264768
1377	86964242 (2753, 2754)	Novel Protein sim. GBank gij1663848 (U75321) - chromaffin granule ATPase II homolog [Mus musculus]	Contains protein domain (PF00122) - E1-E2 ATPase	ATPase-associated	29331824, 264591, 265019, 264686, 264768, 55811957, 264683, 22279002

1378	87595071 (2755, 2756)	Novel Protein sim. GBank g 4107015 db BAA36293 - (AB001772) PEM-5 [Ciona savignyi]				22278994, 22278998, 264093, 264094, 264259, 29331824, 29331827, 285009, 285018, 285019, 18108351, 264369, 264288, 29148627, 55811957, 264891, 18108388, 33857109, 18108388, 264835, 283981, 18108385, 265020
1379	85879344 (2757, 2758)	Novel Protein sim. GBank g 3252872 (AF035620) - BRCA1-associated protein 2 [Homo sapiens]				
1380	87627962 (2759, 2760)	Novel Protein sim. GBank g 4837737 gb AAD30662.1 - (AF098834) germ cell specific Y-box binding protein [Homo sapiens]	nuc_rept			284510, 284512, 285009, 264288, 264564
1381	88179658 (2761, 2762)	Novel Protein sim. GBank g 4731580 gb AAD28508.1 AF12538 - (AF125384) L82A [Drosophila melanogaster]	UNCLASSIFIED			87188559, 265017, 264628, 22279002
1382	94847576 (2763, 2764)			Contains protein domain (PF00583) - Acetyltransferase (GNAT) family		22278994, 22278997, 22278998, 29331822, 29331824, 29331828, 285007, 80432229, 60433396, 85658542, 265017, 285018, 264685, 284768, 21906766, 35695917, 33657023, 27486261, 27488282, 35895763, 35695855, 87168518, 22279002
1383	87860598 (2765, 2766)		UNCLASSIFIED			18108396, 264692
1384	86815895 (2767, 2768)		UNCLASSIFIED			264488, 264508, 264509, 264905, 264906, 264908, 264909, 264511, 264512, 264910, 264760, 18108351, 264768, 264769, 35695855, 264830, 264838, 264555, 264638, 264483, 264584, 264486
1385	86378768 (2769, 2770)	Novel Protein sim. GBank g 2384732 (AF015911) - NAC-1 protein [Rattus norvegicus]	UNCLASSIFIED			35696052, 55811386, 264688, 21906765, 265020, 33657023, 18108385
1386	61013049 (2771, 2772)	Novel Protein sim. GBank g 2384910 (AF022982) - contains similarity to the A-type potassium current class of channel proteins [Caenorhabditis elegans]	Int			60432289, 29331828, 264908, 264907, 58182435, 265011, 264681, 60170815, 33857023, 83373044, 264568
1387	87797958 (2773, 2774)	Novel Protein sim. GBank g 4160304 emb CAA10800 - (AJ132192) HS1 binding protein 3 [Mus musculus]	UNCLASSIFIED			264591
1388	95101652 (2775, 2776)	Novel Protein sim. GBank g 4895184 gb AAD32753.1 AC007231 putative disease resistance protein [Arabidopsis thaliana]	glycoprotein			65274572, 22278999, 284259, 29331826, 29331827, 35896052, 264509, 284907, 264908, 264909, 265008, 285008, 60170831, 33857402, 80433438, 284598, 21908754, 87168559, 264600, 265017, 284883, 18108384, 52844229, 21908785, 21906766, 21906787, 21906788, 21906789, 285021, 284892, 33657023, 33857109, 35895855, 264558, 60170394, 83373044, 22279000

1389	91256016 (2777, 2778)	Novel Protein sim. GBank gll5689387[db][BAA82977.1] - (AB028948) KIAA1025 protein [Homo sapiens]	Contains protein domain (PF00541) - Zn-finger in Ran binding protein and others.	UNCLASSIFIED	65274572, 22278999, 264259, 29331822, 29331824, 29331825, 60432289, 29331828, 29148499, 264908, 68712502, 55812038, 285017, 285018, 265019, 18106351, 264369, 21906765, 21906768, 21906767, 21906768, 285020, 265021, 264692, 33657023, 33657349, 18108370, 18108374, 55811576, 264555, 264556, 284537, 60170394, 83373044, 22279000, 264563, 264564, 52645156, 52648365, 264259, 52645080, 29331825, 29331828, 284908, 52844045, 265009, 33657084, 52644298, 87166474, 87168559, 265017, 265018, 264760, 264682, 264288, 264688, 264687, 56181562, 52644229, 21908765, 21906769, 35695917, 265020, 265021, 60170615, 52844150, 33657023, 27486281, 27486284, 85274791, 264631, 264555, 52644332, 87188518, 22279000, 264567
1390	84111916 (2779, 2780)	Novel Protein sim. GBank gll3702285 (AC005783) - R33083_1 [Homo sapiens]	peptidase		
1391	91227345 (2761, 2782)	Novel Protein sim. GBank gll1346910[sp]P28650[P]UA1_MOUSE - ADENYLOSUCCLINATE SYNTHETASE, MUSCLE ISOZYME (IMP-ASPARTATE LIGASE)	Contains protein domain (PF00709) - Adenylosuccinate synthetase		29331826, 29331828, 29331830, 284448, 284288, 33657023, 18108385, 264555, 264556, 83373044
1392	84311097 (2783, 2784)	Novel Protein sim. GBank gll726286 (U22394) - mSn3A [Mus musculus]			52646842, 65274572, 22278994, 22278995, 35696288, 56994075, 22278997, 22278998, 22278999, 60432049, 264259, 52845080, 29331822, 29331824, 60432289, 29331826, 29331827, 35696052, 29331828, 33656970, 264907, 52844045, 265006, 265007, 285008, 80431735, 60433356, 52646317, 55811366, 52844298, 265010, 87168559, 285017, 284804, 285018, 265019, 264448, 264288, 264369, 264788, 21906764, 21906787, 35695917, 285020, 265021, 33657109, 52645129, 27486261, 27486282, 27486265, 33657349, 35695783, 18108370, 18108374, 18108376, 55811576, 35696423, 35695855, 264636, 52844332, 18108382, 18108385, 87188518, 80432113, 22279000, 264484, 264566, 18108391
1393	80409472 (2785, 2786)		Contains protein domain (PF00560) - Leucine Rich Repeat	UNCLASSIFIED	264783, 264831
1394	15028819 (2787, 2788)			UNCLASSIFIED	264829
1395	95361471 (2789, 2790)	Novel Protein sim. GBank gll2274845[db][BAA21534] - (C68481) N-WASP [Rattus rattus]		UNCLASSIFIED	265009, 18108381

1398	95383253 (2781, 2792)	Novel Protein sim. GBank gl 2135904 p l j 54610 - pHL ETF1 - human			22276997, 22276998, 264259, 29331625, 60432289, 29331826, 29146498, 29146499, 264907, 264908, 29331630, 264909, 265006, 265007, 265008, 265009, 60433358, 265010, 264602, 265017, 265018, 265019, 18108334, 52844228, 16106358, 21906767, 29148627, 21906766, 21906769, 29148629, 29148784, 265021, 265022, 16106366, 18108374, 56182323, 16108385, 264563, 264587, 35696286, 264907, 66712502, 264510, 35695917, 264692, 264693, 35696423, 264259, 29331822, 29331624, 29331625, 29331827, 35696052, 33656970, 67168474, 265016, 265019, 264682, 264768, 21906767, 265020, 33657023, 27486281, 55611576, 264632, 264639, 63373044, 87166516, 22279002
1397	87631317 (2783, 2794)			UNCLASSIFIED	
1398	91233667 (2795, 2798)	Novel Protein sim. GBank gl 5420389 emb CAB46680.1 - (AJ243460) proteophosphoglycan [Leishmania major]			
1399	87631076 (2787, 2798)	Novel Protein sim. GBank gl 2496867 sp Q09232 YQ22_CAEEL - HYPOTHETICAL 32.0 KD PROTEIN C09F5.2 IN CHROMOSOME III		UNCLASSIFIED	
1400	95419084 (2789, 2800)	Novel Protein sim. GBank gl 283920 p l j S27939 - tensin - chicken	Contains protein domain (PF00017) - Src homology domain 2	UNCLASSIFIED	56182575, 22276994, 22278997, 264259, 29331822, 29331825, 29331826, 29331827, 29331828, 264906, 58162435, 264112, 265009, 265011, 265017, 265016, 265019, 264760, 264762, 264765, 294288, 264685, 264687, 56161562, 264769, 21908768, 21908767, 55811957, 264891, 264692, 264628, 264629, 55611576, 264634, 264555, 264637, 264557, 264636, 18108361, 264558, 18108384, 60432113, 22279000 65274572, 60432289, 264909, 264756, 264768, 21906769, 22279002 264905, 264907, 264908, 264909, 264112, 264693, 33657109, 264634
1401	91226379 (2801, 2802)	Novel Protein sim. GBank gl 3256165 emb CAA15485 - (AL008635) dJ510H16.1 [Homo sapiens]	Contains protein domain (PF00780) - VHS domain	UNCLASSIFIED	
1402	95361475 (2803, 2804)	Novel Protein sim. GBank gl 1515427 (U57523) - nel homolog [Homo sapiens]	Contains protein domain (PF00008) - EGF-like domain	igf	
1403	94147933 (2805, 2806)	Novel Protein sim. GBank gl 5262815 emb CAB45747.1 - (AL080156) hypothetical protein [Homo sapiens]			65274572, 66712502, 265017, 264446, 264268, 21906765, 21906769, 264693, 55811578, 65274791, 60432113
1404	90935393 (2807, 2808)			UNCLASSIFIED	65274572, 22278998, 29331622, 29331828, 66712502, 265006, 60433438, 265017, 264693, 18108385

1405	95095068 (2809, 2810)	Novel Protein sim. GBank gij[854065]emb[CAA58337] - (X63413) U88 [Human herpesvirus 6]			264488, 5694075, 35696286, 29331822, 29331824, 29331826, 29331828, 35698052, 264508, 264908, 264907, 264908, 264510, 264511, 264910, 33657402, 264594, 264758, 264600, 264604, 264762, 18106351, 264784, 33657023, 33657109, 284826, 264634, 83373044, 22279002, 264583, 264482, 284488, 264587, 264907, 264605
1406	67812369 (2811, 2812)	Novel Protein sim. GBank gij[624076]b[AAC96425.1] - (U42580) contains Pro-rich Px motifs: SPKPP (20X), PEPPA (9X); similar to soybean pro-rich cell wall protein, corresponds to Swiss-Prot Accession Number P13993 [Parametium bursaria Chlorella virus 1]	collagen		
1407	94129872 (2813, 2814)	Novel Protein sim. GBank gij[2827866] (AF015037) - endodoligopeptidase A related protein; EOPA related protein [Oryctolagus cuniculus]	UNCLASSIFIED		35696286, 22278999, 264094, 264259, 86714117, 29331828, 29331827, 29331828, 29146488, 284107, 264908, 285008, 265008, 264910, 60433438, 265011, 265017, 18108351, 264448, 264288, 264886, 21906765, 21906769, 264692, 33657109, 18106370, 264628, 263972, 18108374, 35696423, 55811576, 264631, 264557, 264558, 63373044, 18108385, 87168518, 60432113, 22279002
1408	95381477 (2815, 2816)	Novel Protein sim. GBank gij[2564953] (AF030001) - unknown [Mus musculus]	Contains protein domain (PF00008) - oncogene EGF-like domain		264488, 264489, 35696286, 264109, 264508, 264905, 264509, 264906, 264907, 264908, 264909, 285008, 265009, 264910, 33657402, 264757, 264758, 285011, 264601, 265017, 264760, 264782, 264683, 264685, 264766, 264687, 264889, 21908767, 265021, 264690, 264691, 33657023, 264692, 264693, 33657109, 264628, 264629, 35696423, 35895855, 264631, 264632, 264634, 264635, 264555, 264638, 264637, 264638, 56182323, 264639, 264563, 264564, 264565, 264568, 264567, 264693
1409	86844385 (2817, 2818)	Novel Protein sim. GBank gij[2662165]dbj[BAA23714] - (AB007902) HH0712 cDNA clone for KIAA0442 has a 574-bp insertion at position 1474 of the sequence of KIAA0442. [Homo sapiens]			
1410	86812597 (2819, 2820)	Novel Protein sim. GBank gij[2493790]sp[Q80994]ACR3_MOUSE - 30 KD ADIPOCYTE COMPLEMENT-RELATED PROTEIN PRECURSOR (ACRP30) (ADIPOCYTE SPECIFIC PROTEIN ADIPOQ)	Contains protein domain (PF00386) - complement C1q domain		29331826, 264112, 264512, 265009, 265010, 264801, 264888, 264789, 21906787, 283974, 264631, 264566

1411	67618641 (2821, 2822)	Novel Protein sim. GBank g 3123155 sp P91343 YIM3M_CAEEL - HYPOTHETICAL 49.0 KD TRP-ASP REPEATS CONTAINING PROTEIN F55F8.5 IN CHROMOSOME I	Contains protein domain (PF00400) - WD domain, G-beta repeat	kinesin receptor	22276995, 22276997, 22276999, 29331622, 60432289, 29331828, 284907, 265017, 265018, 264682, 21906787, 21906766, 21906789, 285020, 284690, 284681, 33657023, 33657108, 27466264, 264626, 263972, 264634, 264558, 16106385 264757
1412	84390819 (2823, 2824)	Novel Protein sim. GBank g 3679121 emb CAA84370 - (Z70310) predicted using GeneFinder; Similarity to Mouse ankyrin (PIR Acc. No. S37771); cDNA EST EMBL:TD1923 comes from this gene; cDNA EST EMBL:D32335 comes from this gene; cDNA EST EMBL:D32723 comes from this gene; cDNA EST E...	Contains protein domain (PF00023) - Ank repeat	UNCLASSIFIED homeobox	56994075, 29331622, 35696052, 29331628, 29331630, 284909, 52844045, 284510, 52844296, 65658542, 67168474, 285017, 265018, 264681, 264687, 21906766, 35695917, 285020, 52644150, 284692, 263967, 27486264, 35695763, 264639, 16100387, 264568 264662, 264663, 265022, 264636
1414	94675860 (2827, 2828)	Novel Protein sim. GBank g 3252961 (AF066892) - Ras- binding protein SUR-8 [Mus musculus]	Contains protein domain (PF00560) - Leucine Rich Repeat	UNCLASSIFIED	52646365, 56162575, 22276894, 22276995, 56994075, 22276996, 22276997, 22276998, 22276999, 264259, 29331622, 29331624, 29331625, 29331826, 29331827, 29331828, 29146496, 66712502, 29331830, 52844045, 264113, 264511, 33657402, 284757, 21906754, 55611388, 265017, 265016, 265019, 264761, 284683, 264369, 264288, 264666, 264689, 21906766, 21906787, 29146827, 21906769, 55611957, 285020, 285021, 284690, 33657023, 65274620, 52645128, 27486282, 27486284, 60431526, 264629, 35695655, 56162323, 284559, 60432113, 264404, 22279002, 264482
1415	94326946 (2828, 2830)	Novel Protein sim. GBank g 1671167 (U90439) - unknown protein [Arabidopsis thaliana]			

1416	94325977 (2831, 2832)	Novel Protein sim. GBank gij5106557[gblAAD39749.1]AF12305 - (AF123052) MLL sepiin-like fusion protein [Homo sapiens]	Contains protein domain (PF00735) - struct Cell division protein	18108392, 18108394, 18108397, 18108398, 22278995, 22278998, 35698286, 22278997, 22278998, 22278999, 29331822, 35696052, 29331826, 29146498, 264905, 264907, 284908, 284828, 284909, 264113, 265006, 285007, 285008, 285009, 60170831, 264595, 18108346, 21906754, 87186474, 265010, 285011, 87188559, 265017, 265018, 265019, 264762, 18108351, 264681, 264783, 264882, 264683, 264788, 52644229, 284886, 284889, 21906785, 21906786, 21906787, 29148827, 21906766, 55811957, 29148628, 265020, 52844150, 18108391, 33857023, 18108382, 18108368, 284828, 18108370, 264629, 18108374, 18108379, 55811578, 65274791, 264634, 284636, 58182323, 18108381, 60170394, 18108385, 58528486, 67188518, 22279000
1417	87826663 (2833, 2834)	Novel Protein sim. GBank gij4958935[dbjBAA78095.1] - (AB027570) suppressor of potassium transport defect 3 [Rattus norvegicus]	ATPase-associated	264107, 264448
1418	87594276 (2835, 2836)		UNCLASSIFIED	264259, 264908, 265010, 52644229, 21908764, 21908788, 264890, 264839, 18108388
1419	87757168 (2837, 2838)	Novel Protein sim. GBank gij2072294 (U85097) - mitotic phosphoprotein 43 [Xenopus laevis]	struct	264259, 60432289, 265006, 87168474, 264286
1420	87286628 (2839, 2840)	Novel Protein sim. GBank gij5174421[refINP_006023.1]pCPNE - copine VI (neuronal) (Z69835) Similarity to Yeast uridine kinase	ATPase-associated	29331824, 265007, 264563
1421	94746986 (2841, 2842)	Novel Protein sim. GBank gij3876090[embjCAA93459.1] - (SW:URK1_YEAST); cDNA EST EMBL:Z14895 comes from this gene; cDNA EST CEMSE17F comes from this gene; cDNA EST EMBL:D67355 comes from this gene; cDNA EST yk209h1.5 comes from this gene...	kinase	18108358, 18108396, 18108397, 21906768, 18108398, 21908787, 58182575, 21908788, 21906769, 56181888, 55811857, 35695917, 35696286, 22278996, 22278997, 22278998, 22278999, 265021, 265022, 60170615, 264259, 33657023, 29331822, 58182181, 29331824, 68714117, 29331825, 33657109, 29331828, 27486281, 29331828, 35696052, 33657349, 264905, 264509, 20281149, 18108370, 264907, 60431528, 88712502, 283972, 55811578, 35696423, 35695855, 264512, 285007, 60431850, 60432229, 60431735, 58182323, 264558, 60170394, 83373044, 55812036, 284758, 18108385, 21906754, 55811386, 87168518, 87168559, 60432113, 285017, 265018, 285019, 22279002, 55811150, 264563, 264682, 284783, 264448, 264586, 284486, 18108391

1422	88178777 (2843, 2844)	Novel Protein sim. GBank gll4505939refNP_000928.1pPOLR - polymerase (RNA) II (DNA directed) polypeptide A (220kD)		mapolymerase	58994075, 35696286, 87168559, 55811957, 55811578, 264555, 264557, 87168518
1423	86997762 (2845, 2846)			UNCLASSIFIED	264686, 264489, 264682, 264594, 264603, 285018, 284908
1424	95201610 (2847, 2848)	Novel Protein sim. GBank gll437181 (U02289) - GTPase-activating protein [Caenorhabditis elegans]	Contains protein domain (PF00820) - sinuc RhoGAP domain		29331822, 29331825, 29331827, 29146498, 264905, 264908, 264908, 264909, 265007, 264910, 265009, 33109954, 285010, 87188559, 265019, 264766, 264687, 21908785, 21908786, 21908787, 21908788, 29148627, 55811957, 29148629, 265021, 264691, 264892, 58526488, 22279002, 264593
1425	21662314 (2849, 2850)	Novel Protein sim. GBank gll100799iprj[S14959 - proline-rich protein - wheat]		UNCLASSIFIED	265007, 264558
1428	94322115 (2851, 2852)	Novel Protein sim. GBank gll2078441 (U56964) - weak similarity to S. cerevisiae intracellular protein transport protein US1 (SP:P25386) [Caenorhabditis elegans]		UNCLASSIFIED	264488, 60424179, 35696288, 22278997, 22278998, 22278999, 264259, 60432049, 29331822, 29331825, 29331827, 59182435, 264910, 60433356, 60433438, 21908754, 265018, 284288, 21908785, 21908786, 21908787, 21908788, 285020, 265022, 33657109, 18108370, 18108378, 264558, 83373044, 18108385, 58526488, 22279002, 264482
1427	91227510 (2853, 2854)	Novel Protein sim. GBank gll5816074jgAA045818.1JAF08194 - (AF061943) prolactin-derived STE20-like kinase PSK [Homo sapiens]	Contains protein domain (PF00069) - kinase Eukaryotic protein kinase domain		22278994, 58994075, 22278997, 29331828, 29331830, 264828, 265008, 265007, 265008, 265009, 264883, 264288, 18108354, 21908785, 21908788, 29148629, 33657023, 18108374, 35895855, 83373044, 22279002, 264584
1428	94323008 (2855, 2856)	Novel Protein sim. GBank gll138350spP28968jVGLX_HSVB - GLYCOPROTEIN X PRECURSOR		glycoprotein	56181688, 284259, 264907, 265007, 265009, 264595, 265010, 264888, 85274620, 284629, 65274791, 22279002, 284588
1429	87888689 (2857, 2858)			UNCLASSIFIED	264112, 264555, 285017, 265019, 21908785, 283977, 284555
1430	94735021 (2859, 2860)	Novel Protein sim. GBank gll1181819dbjBAA11565] - (D82364) a variant of TSC-22 [Gallus gallus]			264094, 29331824, 264591, 264593, 265018, 264681, 21908785, 21908787, 85274820, 55811576, 264839, 87188518, 22279002
1431	80429081 (2861, 2862)	Novel Protein sim. GBank gll5420309fembjCAB46880.1] - (AJ243460) proleophosphoglycan [Leishmania major]		UNCLASSIFIED	284788, 285020, 264891, 264693, 264638
1432	87463004 (2863, 2864)	Novel Protein sim. GBank gll414797 (L18966) - pyruvate dehydrogenase phosphatase [Bos taurus]		phosphatase	18108394, 29146498, 265007, 80433438, 284763, 29148829, 263969
1433	87605403 (2865, 2866)	Novel Protein sim. GBank gll2460318 (AF022147) - uterus-ovary specific putative transmembrane protein [Rattus norvegicus]	Contains protein domain (PF00100) - Zona pellucida-like domain	UNCLASSIFIED	284259, 284510, 284591, 264803, 264565
1434	85713730 (2867, 2868)			UNCLASSIFIED	264882, 264891

1435	94708213 (2869, 2870)	Novel Protein sim. GBank gij3970850[dbj]BAA34789.1] - (AB015330) HRIHF82007 [Homo sapiens]		transcript factor	22278997, 22278998, 22278999, 60432049, 264259, 29331822, 29331824, 66714117, 29331826, 33856970, 264508, 264905, 66712502, 29331830, 264809, 265007, 265008, 264910, 265009, 60433356, 60433438, 264598, 21908754, 265010, 265017, 265018, 265019, 18108351, 264782, 264446, 264288, 264789, 21908787, 21908768, 21908769, 265020, 265021, 265022, 264690, 264691, 33857109, 264828, 18108374, 18108376, 55811576, 264636, 60170394, 56182323, 264559, 83373044, 67168516, 60432113, 22278000, 22278002, 264563, 264482, 264565
1436	86635024 (2671, 2872)	Novel Protein sim. GBank gij3183977[embj]CAA39515] - (X56044) protein HIFC [Mus musculus]		UNCLASSIFIED	263976, 284557, 264559
1437	87631082 (2873, 2874)	Novel Protein sim. GBank gij2496887[spj]Q09232[YQ22_CAEEL - HYPOTHETICAL 32.0 KD PROTEIN CO9F5.2 IN CHROMOSOME III R31240. 1 [Homo sapiens]		UNCLASSIFIED	22278997, 66714117, 29331826, 264907, 56182435, 265008, 18108351, 264692, 264693
1438	85544280 (2875, 2876)	Novel Protein sim. GBank gij1905908 (AD000082) - hypothetical human serine-threonine protein kinase	Contains protein domain (PF00585) - kinase PDZ domain (Also known as DHR or GLGF)		264488, 264508, 264908, 264909, 264757, 264600, 264601, 264805, 264788, 264769, 264890, 35698423, 264556, 264563, 264568
1439	91231694 (2877, 2878)	Novel Protein sim. GBank gij387629[embj]CAA94892] - (Z71180) similar to BPTIKUNITZ inhibitor domain; cDNA EST EMBL:D68293 comes from this gene; cDNA EST yk448h4.5 comes from this gene; cDNA EST yk248e6.5 comes from this gene; cDNA EST yk448h4.3 comes from this gene [Caenorhabdi...]	Contains protein domain (PF00450) - calthepin Serine carboxypeptidase		264488, 18108394, 65274572, 56182575, 22278994, 22278996, 35698286, 22278997, 22278998, 22278999, 264259, 60432049, 29331822, 29331824, 60432288, 29331826, 29331827, 29331828, 35698052, 33856970, 264107, 264508, 264509, 284907, 68712502, 29331830, 56182435, 264511, 265008, 265007, 265009, 60432229, 60433436, 264995, 55812038, 55811386, 265011, 265017, 265018, 265019, 18108351, 264446, 18108354, 264288, 18108355, 264767, 21908765, 21908766, 21908767, 21908768, 21908769, 55811957, 35695917, 285020, 265021, 265022, 33657109, 18108370, 18108374, 55810764, 35695855, 284634, 264636, 56182323, 83373044, 18108387, 87168518, 80432113, 22278000, 264466
1440	87423643 (2878, 2880)	Novel Protein sim. GBank gij2662165[dbj]BAA23714] - (AB007802) HH0712 cDNA clone for KIAA0442 has a 574-bp insertion at position 1474 of the sequence of KIAA0442. [Homo sapiens]		UNCLASSIFIED	264887, 264259, 264808, 264907, 264908, 264909, 265008, 265010, 265017, 265018, 265019, 18108351, 284389, 265020, 33657023, 33657109, 80431528, 55611576, 264635

1441	95317662 (2881, 2882)	Novel Protein sim. GBank glj483956[emb]CAB11123.2] - (Z88551) predicted using hexExon; MAL3P8.28 (PFC0845c). Hypothetical protein, len: 167 aa; Similarity to model organism hypothetical proteins (C.elegans, D.melanogaster, S.cerevisiae & S.pombe). C.elegans protein ZK287.5 (TR:...	Contains protein domain (PF00846) - F-box domain.	helicase	18108392, 264488, 263984, 264489, 58182575, 22278984, 22278995, 58994075, 35696288, 22278997, 22278999, 60432049, 264259, 29331822, 29331824, 56182181, 66714117, 29331828, 29331827, 29331828, 35698052, 29148498, 264508, 264905, 264509, 264906, 264907, 264908, 264909, 58182435, 264510, 264511, 264512, 264910, 264592, 264593, 33857402, 80433438, 264595, 264758, 21906754, 85858542, 87188474, 285010, 87168559, 264600, 264602, 265017, 264604, 285018, 264805, 285019, 264760, 264761, 264762, 264881, 284448, 264784, 264683, 264288, 284768, 264768, 264769, 52844228, 264688, 21908785, 21908788, 21908787, 21908788, 21908789, 55811957, 35895917, 265020, 265021, 80170615, 52644150, 264681, 264692, 33657023, 85274820, 33857108, 35695763, 264628, 18108370, 204629, 18108374, 55811576, 35698423, 65274791, 35695855, 264631, 264834, 264635, 264836, 264637, 264838, 56182323, 264558, 60170394, 264638, 264559, 83373044, 18108385, 18108388, 58526488, 22278000, 22279002, 284583, 284483, 264564, 264588, 264587, 264488
1442	63367491 (2883, 2884)	Novel Protein sim. GBank glj5103027[dbj]BAA78765.1] - (AB023419) mSox7 [Mus musculus]		transcript/factor	264906, 285007, 264693, 264558
1443	87109835 (2885, 2886)	Novel Protein sim. GBank glj4887229[gb]AAD32244.1[AF150755] - (AF150755) microtubule-actin crosslinking factor [Mus musculus]	Contains protein domain (PF00435) - Spectrin repeat	stnuc	52645080, 264691, 264628, 264555
1444	87820478 (2887, 2888)	Novel Protein sim. GBank glj3874447[emb]CAB02772] - (Z81039) predicted using GeneFinder; cDNA EST EMBL: T01209 comes from this gene; cDNA EST yk278a11.3 comes from this gene; cDNA EST yk308a9.3 comes from this gene; cDNA EST yk308a9.5 com...		UNCLASSIFIED	264258, 29331822, 29331824, 66714117, 29331828, 284288, 35895917, 33657023, 284835, 60170394
1445	94890470 (2889, 2890)	Novel Protein sim. GBank glj295886[emb]CAA11022] - (AJ222968) L-perilaxin [Mus musculus]		UNCLASSIFIED	264369
1446	85079288 (2891, 2892)	Novel Protein sim. GBank glj5081810[gb]AAD39464.1[AF135440] - (AF135440) huntington yeast partner C [Mus musculus]	Contains protein domain (PF01848) - FF domain		264369
1447	86845392 (2893, 2894)				16106398, 35685286, 22278897, 66714117, 29331828, 285009, 264758, 285018, 264288, 21906788, 21808787, 264892, 284834, 264566
1448	94990477 (2895, 2896)	Novel Protein sim. GBank glj3980411 (AC004561) - putative proline-rich protein [Arabidopsis thaliana]	Contains protein domain (PF00439) - Bromodomain		29331827, 284508, 284908, 285008, 264595, 18108357, 18108385, 264568, 284488

1449	87860859 (2897, 2898)					66714117, 284908, 284908, 284591, 284601, 284764, 284832
1450	87458696 (2899, 2900)	Novel Protein sim. GBank gll1707074 (U80450) - MOTIE11.2 [Caenorhabditis elegans]			UNCLASSIFIED	35696286, 35696052, 265008, 265009, 60170831, 33109854, 264883, 284689, 35896423, 35895855, 58528488
1451	87797970 (2801, 2802)	Novel Protein sim. GBank gll180304 (emb) [CAA10600] - [AJ132192] HS1 binding protein 3 [Mus musculus]			UNCLASSIFIED	29331828, 284683, 284893, 283978, 264630
1452	85692898 (2803, 2804)	Novel Protein sim. GBank gll2832806 (db) [BAA24608.1] - [D88340] dipeptidyl peptidase III [Rattus norvegicus]			peptidase	264681, 33657023, 284829
1453	88130434 (2805, 2806)	Novel Protein sim. GBank gll28831 (sp) [P38188] ALU1_HUMAN - IIII ALU SUBFAMILY J WARNING ENTRY IIII			kinase	264510, 264768
1454	11204698 (2807, 2808)					264556
1455	87797898 (2809, 2810)				UNCLASSIFIED	29331822, 68714117, 29331825, 264905, 29331830, 265008, 265008, 285008, 285011, 265019, 18108351, 21908768, 33857109, 18108378, 284632, 58182323, 87168518
1456	86320218 (2811, 2812)	Novel Protein sim. GBank gll29230 (sp) [P41004] CUT3 SCHPO - CHROMOSOME SEGREGATION PROTEIN CUT3			transport	22278995, 22278996, 22278997, 22278998, 22278999, 29331827, 284107, 285017, 21906765, 21906788, 21908787, 21908769, 28148629, 18108370, 22279000
1457	80076800 (2813, 2814)				UNCLASSIFIED	264107, 264568
1458	87800460 (2815, 2816)	Novel Protein sim. GBank gll2246532 (U93872) - ORF 73, contains large complex repeat CR 73 [Kaposi's sarcoma-associated herpesvirus]			transport	56182575, 22278999, 80432049, 284259, 29331826, 29331827, 29331828, 264102, 264107, 264110, 285009, 60432229, 265019, 285020, 283972, 263978, 284835, 22279002, 264566
1459	95360820 (2817, 2818)	Novel Protein sim. GBank gll5524687 (gb) [AAD44333.1] AF15935 - (AF159356) Munc13-4 protein [Rattus norvegicus]	Contains protein domain (PF00168) - C2 domain		kinase	22278997, 284259, 29331824, 29331828, 29331827, 29331828, 285017, 285018, 264760, 264682, 264448, 264288, 264766, 265021, 284892, 33857023, 33857109, 35695855, 284568
1460	95354802 (2819, 2820)				UNCLASSIFIED	29331822, 264591, 55811957, 284691, 284693, 65274820
1461	94741513 (2821, 2822)	Novel Protein sim. GBank gll1707274 (U80931) - strong similarity to class-III of pyridoxal-phosphate-dependent aminotransferases [Caenorhabditis elegans]	Contains protein domain (PF00202) - gaba Aminotransferases class-III pyridoxal phosphate			22278997, 29331822, 35696052, 265009, 264758, 265017, 285018, 285019, 284760, 264389, 284887, 21908765, 21908788, 265022, 33657109, 27486261, 264555, 83373044
1462	87732016 (2823, 2824)				UNCLASSIFIED	264555, 264558
1463	88090605 (2825, 2826)	Novel Protein sim. GBank gll1770466 (emb) [CAA68912] - (X98259) M-phase phosphoprotein 8 [Homo sapiens]	Contains protein domain (PF00385) - 'chrome' (CHR) domain Organization (Mofifer) domain			80432049, 264259, 29146499, 284906, 264907, 284512, 285017, 284783, 264768, 18108370, 18108374, 284838, 18108385, 18108388

1464	87820482 (2827, 2928)	Novel Protein sim. GBank gl 3874447 emb CA802772 - (Z81039) predicted using GeneFinder: cDNA EST EMBL:U01208 comes from this gene; cDNA EST yk278a11.3 comes from this gene; cDNA EST yk308a9.3 comes from this gene; cDNA EST yk308a9.5 com...		UNCLASSIFIED	264569, 22278995, 22278998, 22278997, 22278996, 29331822, 29331824, 29331825, 35696052, 20281100, 264905, 29331830, 264909, 265007, 33857402, 21908754, 265017, 285018, 264682, 264684, 264369, 264288, 264766, 21908765, 21908788, 21908767, 21906769, 35895917, 264691, 33657023, 264692, 35696423, 35695955, 264630, 284831, 264839, 264565
1465	87425192 (2829, 2930)	Novel Protein sim. GBank gl 4569599 gb BAA76821.1 - (AB023194) KIAA0877 protein [Homo sapiens]		glucoamy/lase	264488, 22278994, 56954075, 60432049, 264259, 58182181, 80432289, 29331827, 52844045, 284511, 285007, 285008, 284596, 55812038, 55811388, 284600, 264602, 285017, 285018, 264604, 285019, 18108351, 18108354, 58181562, 21908789, 285021, 33657023, 33857162, 55811578, 264557, 18108362, 60432113
1466	87606227 (2931, 2932)	Novel Protein sim. GBank gl 2248532 (U93872) - ORF 73, contains large complex repeat CR 73 [Kaposi's sarcoma-associated herpesvirus]		UNCLASSIFIED	264512, 265017, 264689, 264558
1467	87614328 (2933, 2934)	Novel Protein sim. GBank gl 4507241 ref NP_003137.1 pSSRP - structure specific recognition protein 1	struct		264683, 284636
1468	95342862 (2935, 2936)	Novel Protein sim. GBank gl 4507241 ref NP_003137.1 pSSRP - structure specific recognition protein 1			
1469	79236174 (2937, 2938)	Novel Protein sim. GBank gl 1906596 (U81786) - kinesin-73 [Drosophila melanogaster]	struct		22278998, 284758, 285018, 285019, 21908769, 265020, 33857109, 22279002
1470	84990482 (2939, 2940)	Novel Protein sim. GBank gl 5649170 gb AAD43131.2 AF15909 - (AF159092) syd709613 protein [Homo sapiens]		UNCLASSIFIED	18108394, 18108397, 18108398, 35696052, 29148499, 285007, 265008, 285009, 285010, 285011, 18108354, 18108385, 18108386, 18108374, 18108381, 18108382, 18108384, 18108388
1471	87828842 (2941, 2942)	Novel Protein sim. GBank gl 3878146 emb CA801750 - (Z78542) similar to Mitochondrial carrier proteins; cDNA EST EMBL:U01651 comes from this gene [Caenorhabditis elegans]	transport		18108394, 18108398, 56182575, 264259, 29331822, 29331824, 29331825, 80432289, 284907, 284909, 285007, 284910, 285009, 284591, 80432229, 60433358, 264595, 60433438, 284758, 33109954, 265010, 285011, 265018, 264780, 284448, 284764, 264288, 284368, 18108357, 264769, 18108358, 21908767, 21908789, 55811957, 285021, 18108361, 284891, 18108362, 18108365, 18108368, 284628, 18108378, 284837, 284557, 18108381, 56182323, 18108382, 83373044, 18108364, 18108388, 87168518, 60432113, 264404, 22279002, 284492, 264587, 264487
					29331822, 29331824, 29331825, 264828, 284603, 284698, 264893, 18108374, 55811578

1472	67756616 (2843, 2844)	Novel Protein sim. GBank gi 4660707 gb AAD27743.1 AF13296 - (AF13296) CGI-34 protein [Homo sapiens]		UNCLASSIFIED	264905, 16106351, 21906765, 264486
1473	67791809 (2845, 2846)	Novel Protein sim. GBank gi 3686780 (AF042180) - testis- specific Y-encoded-like protein [Mus musculus]	Contains protein domain (PF00956) - MHC Nucleosome assembly protein (NAP)		16106394, 22276895, 56994075, 22276899, 29331622, 29331624, 66714117, 29331825, 29331826, 35696052, 284908, 284907, 56162435, 285007, 284756, 285016, 285019, 284760, 284764, 284288, 284665, 284666, 284766, 21906768, 55611957, 285021, 284691, 284693, 284629, 55611576, 284634, 284636, 56162323, 22279002, 284566, 284466
1474	65600989 (2847, 2848)	Novel Protein sim. GBank gi 2494600 sp Q92176 CORO_BOVIN - CORONIN-LIKE PROTEIN P57	Contains protein domain (PF00400) - struct WD domain, G-beta repeat		284466, 35695917, 35696286, 264692, 33657023, 284693, 33657109, 35696052, 284508, 284905, 284908, 284907, 284629, 284906, 284909, 35696423, 35695655, 284511, 284910, 284632, 284634, 284635, 284636, 284637, 284556, 284557, 284639, 284756, 60432113, 284604, 284605, 284565, 284566, 284764, 284466, 284665, 284766, 284681, 284682, 264286, 284566
1475	86671935 (2849, 2850)		Contains protein domain (PF00041) - UNCLASSIFIED Fibronectin type III domain		
1476	67548855 (2851, 2852)	Novel Protein sim. GBank gi 475752 ref NP_004664.1 pANGP - angiotensinogen 3	Contains protein domain (PF00147) - glycoprotein Fibrinogen beta and gamma chains, C-terminal globular domain		60424178, 56161686, 29331624, 60424289, 29331826, 35696052, 284506, 284905, 284906, 284907, 284908, 284909, 284512, 285007, 285008, 285009, 284910, 33657402, 284595, 284596, 55612036, 285011, 284601, 284762, 16106351, 284286, 284369, 284685, 284766, 284689, 55611957, 284891, 284692, 284693, 16106370, 60431526, 16106374, 35696423, 284634, 284635, 284636, 60431850, 264555, 284638, 284557, 264639, 16106362, 16106366, 60432113, 22279002
1477	67774279 (2853, 2854)	Novel Protein sim. GBank gi 2496306 sp Q80670 DP1_MOUSE - POLYPOSIS LOCUS PROTEIN 1 HOMOLOG (TB2 PROTEIN HOMOLOG) (GP106)		UNCLASSIFIED	264259, 264107, 284905, 265008, 265010, 285011, 284682, 284288, 285020, 285021, 263974
1478	11754412 (2855, 2856)				264686

1479	91640140 (2857, 2858)	Novel Protein sim. GBank gi 5495741 gb AAD3978.1 AF15298 - (AF15298) chromatin-specific transcription elongation factor FACT 140 kDa subunit [Homo sapiens]		peptidase	56182575, 22278995, 22278996, 22278998, 22278999, 29331822, 29331824, 66714117, 264908, 264907, 56182435, 285008, 60170831, 33857402, 284758, 33109854, 21908754, 265017, 265016, 284448, 284288, 264767, 264667, 52644229, 21906764, 264689, 21908785, 21908788, 21908789, 265020, 265021, 60170615, 264691, 33857023, 33857109, 33857182, 27486261, 27486282, 33857348, 18108370, 60431528, 263976, 55811576, 264556, 264557, 60170394, 87166518, 284404, 22279000, 22279002, 264563, 264482
1480	94312412 (2959, 2960)	Novel Protein sim. GBank gi 3550456 emb CAA06329.1 - (AJ005073) Alix [Mus musculus]		UNCLASSIFIED	18108394, 65274572, 56183575, 22278995, 35698286, 56994075, 22278996, 22278997, 22278998, 22278999, 284091, 264259, 35698052, 29148499, 284103, 284105, 284108, 284907, 52844045, 284112, 265007, 265008, 265009, 60433356, 60433436, 264598, 33109934, 33857084, 52844286, 87166474, 265010, 87168559, 285017, 265018, 265019, 264448, 264682, 264683, 264768, 21908785, 21908786, 21908787, 21908788, 21908789, 265020, 265021, 60170615, 52844150, 33857109, 33857182, 263972, 35695855, 264557, 263981, 83373044, 18108385, 87168518, 264566, 264089, 264486
1481	67021442 (2961, 2962)	Novel Protein sim. GBank gi 4836807 gb AAD30568.1 AF14679 - (AF146793) PFT27 [Mus musculus]		MHC	265006, 265007, 265010, 18108374
1482	65320442 (2863, 2864)	Novel Protein sim. GBank gi 4565372 gb AAD25403.1 AF12282 - (AF122923) Vht inhibitory factor-1 [Mus musculus]	Contains protein domain (PF00008) - EGF-like domain	UNCLASSIFIED	264908, 264910, 264755
1483	94115503 (2865, 2866)	Novel Protein sim. GBank gi 535426 (U13736) - calmodulin- like protein [Pisum sativum]	Contains protein domain (PF00038) - struct EF hand	struct	264259, 29331822, 52845080, 29331825, 29331826, 33856970, 29331830, 265007, 55812038, 33109854, 265017, 284288, 21908788, 21908789, 284636, 16108380, 87166518, 22279000
1484	94131544 (2967, 2968)	Novel Protein sim. GBank gi 1911774 gb 180090 - (S83364) putative Rab5-interacting protein (clone L1-57) [human, HeLa cells, Peptide Partial, 122 aa] [Homo sapiens]		UNCLASSIFIED	264489, 35698286, 264256, 264107, 264909, 265008, 60433356, 33857402, 60433436, 264288, 21906765, 21908786, 29148627, 33857023, 27486282, 18108374, 35698423, 83373044, 60432113
1485	60194441 (2969, 2970)	Novel Protein sim. GBank gi 5360129 gb AAD42863.1 AF15511 - (AF155117) NY-REN 62 enlign [Homo sapiens]	Contains protein domain (PF00225) - struct Kinesin motor domain	struct	264369, 265020, 18108374

1486	94125066 (2971, 2972)	Novel Protein sim. GBank gll45895181b1b1BAA76780.11 - (A8023153) KIAA0936 protein [Homo sapiens]	Contains protein domain (PF00069) - kinase	56182575, 22278998, 264908, 264907, 21906754, 87168474, 265017, 285019, 18108351, 264288, 265020, 264568, 21906754, 264486
1487	68452711 (2973, 2974)	Novel Protein sim. GBank gll50192751emb1CAB44431.11 - (AJ132751) xenobiotic/medium-chain fatty acid:CoA ligase [form XL-III] [Bos taurus]	synthase	
1488	87732026 (2975, 2978)	Novel Protein sim. GBank gll5712131gb1AAD47379.1JAF12049 - (AF120499) DEM1 protein [Homo sapiens]	Contains protein domain (PF01443) - fgl Viral (Superfamily 1) RNA helicase	264686, 264769, 264888, 264692, 264693, 264509, 264908, 264907, 18108370, 264908, 264629, 264909, 264510, 265008, 264512, 285007, 285008, 285009, 264555, 264556, 264557, 264558, 264762, 264564, 264682, 21906767, 22278999, 285022, 264259, 29331827, 29331824, 29331825, 29331826, 29331827, 29331828, 284103, 263972, 68712502, 35896423, 35895855, 265007, 265008, 265009, 83373044, 21908754, 58526488, 285017, 264583, 18108351, 264564, 264568, 284369, 264288
1489	95104277 (2977, 2978)	Novel Protein sim. GBank gll2497303jsp1Q627861FPRP_RAT - PROSTAGLANDIN F2-ALPHA RECEPTOR REGULATORY PROTEIN PRECURSOR (PROSTAGLANDIN F2-ALPHA RECEPTOR ASSOCIATED PROTEIN)	Contains protein domain (PF00047) - prostaglandin Immunoglobulin domain	56182575, 264259, 29331822, 29331824, 66714117, 29331827, 29331828, 284508, 264905, 68712502, 265007, 265008, 264594, 33657402, 55812038, 87189474, 285018, 18108351, 264369, 264288, 264769, 264689, 21906767, 21906768, 55811957, 60170815, 33657109, 35895855, 264635, 60170394, 56526486, 22279002, 264563, 265007, 264448, 18108372, 264558, 56182323
1490	87380127 (2979, 2980)		UNCLASSIFIED	
1491	83594305 (2961, 2982)	Novel Protein sim. GBank gll295671 (L11275) - selected as a weak suppressor of e mutant of the subunit AC40 of DNA dependant RNA polymerase I and III [Saccharomyces cerevisiae]	UNCLASSIFIED	
1492	85805363 (2963, 2984)	Novel Protein sim. GBank gll1656005 (U71205) - rti [Mus musculus]	Contains protein domain (PF00071) - oncogene Ras family	22278997, 22278998, 28331822, 264907, 68712502

1493	91677215 (2865, 2886)	Novel Protein sim. GBank gi 5689515 dbj BAA83041.1 - (AB028012) KIAA1089 protein [Homo sapiens]		UNCLASSIFIED	264488, 52646365, 65274572, 56182575, 22278994, 35698286, 56994075, 22278999, 60432049, 28331824, 29331828, 35698052, 264508, 264905, 264906, 52844045, 264909, 56162435, 265006, 265008, 265009, 60170831, 33657402, 55812036, 265010, 265011, 265017, 265016, 265019, 55811150, 264448, 264682, 264685, 264686, 52844229, 21906765, 21906766, 21906768, 21906769, 265020, 265021, 80170615, 52644150, 33657023, 16106364, 18106365, 33657109, 33657182, 27486261, 27486262, 27486264, 33657349, 27486265, 35695763, 18108370, 264829, 18108374, 52844332, 56182323, 87168516, 22279002, 264564, 264566, 264567
1494	87805265 (2887, 2888)	Novel Protein sim. GBank gi 728632 sp P39189 ALU2_HUMAN - III ALU SUBFAMILY SB WARNING ENTRY III	Contains protein domain (PF01352) - kinase KRAB box	UNCLASSIFIED	264807, 265009, 264769, 18108370, 55811576, 264639, 264565, 264488
1495	87605267 (2889, 2890)	Novel Protein sim. GBank gi 4589588 dbj BAA76616.1 - (AB023169) KIAA0972 protein [Homo sapiens]	Contains protein domain (PF01352) - transcript factor KRAB box	UNCLASSIFIED	22278997, 264259, 264906, 264907, 265009, 264594, 33657064, 265017, 264760, 264448, 33657109, 264630, 264634, 58526486, 264563, 264565, 264566, 264488, 264567
1496	87764322 (2891, 2892)	Novel Protein sim. GBank gi 5420387 emb CA846679.1 - (AJ243459) proteophosphoglycan [Laishanla major]		UNCLASSIFIED	35698286, 264906, 265019, 264683
1497	81695428 (2893, 2894)	Novel Protein sim. GBank gi 3674925 emb CAA92591.1 - (Z66296) Similarly to Mouse A-RAF proto-oncogene serine/threonine-protein kinase (SW:KRAA_MOUSE); cDNA EST EMBL:ID27610 comes from this gene; cDNA EST EMBL:TO1018 comes from this gene; cDNA EST EMBL:D33256 comes from this gene...	Contains protein domain (PF00089) - kinase Eukaryotic protein kinase domain	UNCLASSIFIED	264910, 264758, 265011, 264764, 264286, 264690, 264634, 264635, 56526488
1498	80834838 (2895, 2896)	Novel Protein sim. GBank gi 728636 sp P39189 ALU6_HUMAN - III ALU SUBFAMILY SP WARNING ENTRY III		oncogene	264468, 65274572, 28331822, 265017, 265018, 21906765, 29148627, 29148629, 18108374, 264637, 264638, 264567
1499	86451589 (2897, 2898)	Novel Protein sim. GBank gi 2570198 (U54556) - microfilament sheath protein SHP3 [Lilomiosoides sigmodontis]		glucoamylase	263878, 264566
1500	80498388 (2899, 3000)	Novel Protein sim. GBank gi 2078483 (U43200) - antifreeze glycoprotein AFGP polypeptide precursor [Boreogadus saida]		UNCLASSIFIED	22278998, 264769, 18108379
1501	85795287 (3001, 3002)			UNCLASSIFIED	264559
1502	80206141 (3003, 3004)				264508, 264112, 264604, 264684, 52644150, 55811576, 264632, 264556, 264638, 56182323, 264563, 264488
1503	67012701 (3005, 3006)	Novel Protein sim. GBank gi 3900655 (AC004674) - similar to N-acetylglucosaminyltransferase; similar to Q07537 (PID:gi1171989) [Homo sapiens]	Contains protein domain (PF00535) - transferase Glycosyl transferases	UNCLASSIFIED	29331822, 265007, 264369

1504	79840051 (3007, 3008)		Contains protein domain (PF00023) - Ank repeat	UNCLASSIFIED	264693	
1505	85102672 (3008, 3010)	Novel Protein sim. GBank gll4753775jmbjCAB41970.1] - (AJ132545) protein kinase [Homo sapiens]	Contains protein domain (PF00069) - Eukaryotic protein kinase domain	kinase	29331828, 35686052, 264509, 264906, 284907, 284908, 284909, 284511, 264910, 55812036, 264759, 264783, 264784, 284889, 35695917, 285022, 33657109, 18108374, 284631, 284635, 284638, 284568	
1508	94143219 (3011, 3012)	Novel Protein sim. GBank gll1304201dbjBAA081701 - (D29768) alternatively spliced product [Rattus norvegicus]	Contains protein domain (PF00018) - SH3 domain	glycoprotein	65274572, 58182575, 56894075, 22278897, 22278898, 22278899, 284091, 264092, 60432049, 284259, 52645080, 28331822, 29331827, 284108, 28331830, 264908, 58182435, 284110, 284511, 284512, 55812036, 21906754, 87168559, 264600, 265017, 265018, 284681, 18108354, 264389, 264887, 284889, 21908765, 29148627, 21906769, 21906769, 29148829, 52844150, 33857023, 18108378, 65274791, 56182323, 284558, 284559, 18108385, 87168518, 60432113, 22279000, 284555	
1507	83738250 (3013, 3014)	Novel Protein sim. GBank gll5689513dbjBAA03040.1] - (AB029011) KIAA1088 protein [Homo sapiens]		helicase	264639	
1508	11818758 (3015, 3016)	Novel Protein sim. GBank gll5031975jrelnp_005875.1lpPAK4 - protein kinase related to S. cerevisiae STE20, effector for Cdc42Hs	Contains protein domain (PF00069) - Eukaryotic protein kinase domain	kinase	264593	
1509	87318451 (3017, 3018)				264259, 29331822, 29331824, 29331828, 264591, 33109954, 284583	
1510	95362843 (3019, 3020)	Novel Protein sim. GBank gll113181lspIP28614JACOR_ALCEU - ACETON CATABOLISM REGULATORY PROTEIN		UNCLASSIFIED	264259, 29331822, 265007, 18108374, 264556	
1511	88318073 (3021, 3022)	Novel Protein sim. GBank gll728631lspIP38188JALU1_HUMAN - III ALU SUBFAMILY J WARNING ENTRY III		UNCLASSIFIED	264259, 29331822, 29331824, 29331827, 264509, 284907, 284510, 284511, 265007, 264512, 285008, 87168559, 284288, 265022, 33657023, 35895855, 284637, 284638, 284583	
1512	95345390 (3023, 3024)	Novel Protein sim. GBank gll4559353jgbJAD23014.1AC00658 - (AC006585) putative extragenic suppressor protein [Arabidopsis thaliana]	Contains protein domain (PF01183) - RIO1/ZK832.3/MLJ0444 family		52645156, 18108396, 56894075, 80432289, 285008, 80433396, 80433438, 21908754, 87188474, 87188559, 285018, 264782, 264783, 284887, 21908785, 21906769, 27488282, 35695783, 18108374, 35896423, 264555, 18108385, 18108387, 18108388, 87188518, 284482	
1513	87436228 (3025, 3026)	Novel Protein sim. GBank gll1330394 (U58761) - C01F1.6 gene product [Caenorhabditis elegans]			35698052, 264905, 264906, 264907, 264908, 264909, 264910, 264591, 284788, 284689, 284692, 284829, 264838	

1514	95345392 (3027, 3028)	Novel Protein sim. GBank gi 4559333 gb AAD23014.1 AC00658 - (AC006585) putative extrogenic suppressor protein [Arabidopsis thaliana]	Contains protein domain (PF01163) - R1017K632.3/MJ0444 family	UNCLASSIFIED	52644507, 52645156, 52646365, 52646842, 85274572, 22278994, 35898286, 58994075, 284259, 52645090, 29331822, 29331825, 35896052, 29331830, 52644045, 56182435, 285006, 60433356, 60433436, 55812038, 21906754, 526448317, 52644298, 87168474, 87168559, 284448, 52644229, 21906785, 21906786, 21908787, 21908766, 35895917, 265020, 52644150, 33857023, 52645129, 33857109, 33857182, 27486261, 27488262, 27486284, 27486285, 35895783, 18108376, 35898423, 35895855, 52644332, 18108385, 18108387, 87168518, 60432113 265020, 264639
1515	78163536 (3029, 3030)	Novel Protein sim. GBank gi 3879501 emb CAA87795 - (Z47812) similar to ubiquitin carboxyl-terminal hydrolase; cDNA EST EMBL:D33368 comes from this gene; cDNA EST EMBL:D33965 comes from this gene; cDNA EST EMBL:D33822 comes from this gene; cDNA EST EMBL:D34547 comes from this ge...	ubiquitin	UNCLASSIFIED	
1516	88073538 (3031, 3032)	Novel Protein sim. GBank gi 498015 (L27479) - X123 [Homo sapiens]	UNCLASSIFIED		
1517	87793325 (3033, 3034)	Novel Protein sim. GBank gi 3415134 (AF082024) - Phytb1 [Pimpinella brachycarpa]			285008, 56182323, 22279002
1518	87350687 (3035, 3036)	Novel Protein sim. GBank gi 728838 sp P39195 ALU8_HUMAN - III ALU SUBFAMILY SX WARNING ENTRY III			264091, 18108370, 264404
1519	84328689 (3037, 3038)	Novel Protein sim. GBank gi 5282661 emb CAB45771.1 - (AL080198) hypothetical protein [Homo sapiens]	Im7		68714117, 264508, 264509, 264805, 264510, 264910, 264591, 264595, 264288, 264788, 264769, 18108374, 284838, 264838, 264486 284569, 284489, 60432049, 265009, 33857402, 264598, 21908754, 265019, 264369, 21908785, 21908768, 21908769, 264691, 65274620, 33657182, 27488281, 18108374, 264557, 264839, 87168518, 22279002
1520	87592855 (3039, 3040)	Novel Protein sim. GBank gi 2662161 db BAA23712 - (AB007900) H40432 cDNA clone for KIAA0440 has a 438- bp insertion at position 1711 of the sequence of KIAA0440. [Homo sapiens]	UNCLASSIFIED		18108392, 60432049, 264259, 29331824, 265007, 60433356, 265010, 21908768, 264838
1521	86970698 (3041, 3042)	Novel Protein sim. GBank gi 5052351 gb AAD38518.1 AF13542 - (AF135421) GDP. mannose pyrophosphorylase B [Homo sapiens]	Contains protein domain (PF00483) - Nucleotidyl transferase	synthase	18108394, 264259, 68714117, 265011, 264603, 265019, 18108384, 35898423, 264557, 264558, 18108388
1522	78960587 (3043, 3044)	Novel Protein sim. GBank gi 3778587 (AC005388) - Strong similarity to F21B7.33 gi 2808264 from A. thaliana BAC gb AC002560. EST gb N65119 comes from this gene. [Arabidopsis thaliana]	UNCLASSIFIED		29331824, 265018, 285020, 285021
1523	91005151 (3045, 3046)		UNCLASSIFIED		65274572, 21908768, 264883
1524	80203723 (3047, 3048)				
1525	87799867 (3049, 3050)	Novel Protein sim. GBank gi 4759040 ref NP_004283.1 pRIN1 - ras inhibitor	UNCLASSIFIED		284112, 21908754, 283974 284883, 284887, 284686, 264690, 264692, 284883

1528	95105344 (3051, 3052)	Novel Protein sim. GBank gi728850 sp P08640 AMYH_YEAST - GLUCOAMYLASE S1/S2 PRECURSOR (GLUCAN 1,4-ALPHA- GLUCOSIDASE) (1,4-ALPHA-D-GLUCAN GLUCOHYDROLASE)			35698286, 58182181, 60431735, 264595, 55812038, 284805, 284883, 21908785, 55811857, 285020, 65274791, 264555, 264556, 264557, 284558, 264559, 83373044
1527	88282512 (3053, 3054)	Novel Protein sim. GBank gi7292406 (AF041107) - tulip 2 [Rattus norvegicus]			56182575, 264259, 60432049, 29331822, 60432289, 284908, 86712502, 60433438, 87168559, 285017, 264288, 21908768, 21906769, 283977, 55811576, 56182323, 18108381
1528	94130918 (3055, 3058)			UNCLASSIFIED	22278895, 22278897, 264259, 66712502, 264598, 285017, 285018, 284882, 284448, 264683, 264784, 264685, 284686, 21908785, 21908766, 21908787, 21908768, 21908769, 265022, 264693, 83373044, 18108385
1529	94120793 (3057, 3058)	Novel Protein sim. GBank gi4406683 gb AAD20053 - (AF131828) Unknown [Homo sapiens]		UNCLASSIFIED	284488, 263994, 58182575, 22278895, 35698288, 22278897, 264259, 29331822, 60432289, 29331827, 35698052, 264509, 284908, 284907, 284908, 284909, 52644045, 58182435, 284511, 285009, 284910, 60433358, 60433438, 285017, 285018, 264760, 264448, 264764, 264389, 264288, 264768, 18108357, 264788, 52644229, 21908765, 21906768, 21906787, 21908768, 265021, 285022, 52644150, 33657109, 264628, 35695855, 80432113, 22279002, 264563, 284584, 264488, 264567
1530	95012785 (3059, 3060)	Novel Protein sim. GBank gi2828710 (AF043642) - matrix cytophlin [Rattus norvegicus]			264488, 264489, 35696288, 29331825, 35696052, 284508, 284905, 264908, 284907, 264909, 264510, 264511, 284512, 284910, 264592, 284595, 18108351, 284784, 284883, 264684, 284768, 264768, 18108357, 264789, 35695917, 284828, 284828, 18108374, 35695855, 264630, 284831, 264634, 264555, 264638, 284837, 264404, 264583, 264588, 264488
1531	95419351 (3061, 3062)	Novel Protein sim. GBank gi1805874 (U90878) - carboxyl terminal LIM domain protein [Homo sapiens]	Contains protein domain (PF00595) - PDZ domain (Also known as DHR or GLGF).	kinase	56182575, 35696288, 284097, 264259, 29331822, 29331825, 29331828, 29331827, 35696052, 284509, 58182435, 264510, 264511, 265007, 80433356, 55811388, 264681, 284359, 264288, 284766, 264687, 55811957, 35695917, 33657023, 35695783, 55810764, 35696423, 55811576, 283981, 60170394, 56182323, 83373044, 60432113, 264568

1532	85718224 (3063, 3064)	Novel Protein sim. GBank gi 3674716 emb CAA91265 (Z86494) cDNA EST EMBL:D63271 comes from this gene; cDNA EST EMBL:D64845 comes from this gene; cDNA EST EMBL:D64449 comes from this gene; cDNA EST EMBL:D67438 comes from this gene; cDNA EST EMBL:D68087 comes from this gene; cDNA...			UNCLASSIFIED	264689
1533	94239830 (3065, 3068)	Novel Protein sim. GBank gi 1490324 emb CA801543 (Z78141) unknown [Mus musculus]			slud	29331824, 29146498, 284607, 284112, 265008, 285011, 285017, 285018, 284782, 18108351, 283867, 20281148, 18108374, 263981, 264568
1534	95343841 (3067, 3068)	Novel Protein sim. GBank gi 81286 pir IS2687 - extensin - Volvox carter (fragment)			UNCLASSIFIED	264905, 264907, 264788, 264637
1535	90393732 (3068, 3070)					65274572, 22278997, 264259, 60432049, 29331822, 80432288, 29331827, 28146499, 265006, 285008, 60170831, 80433438, 33109954, 87188559, 265018, 18108357, 21908788, 28146628, 265021, 265022, 18108377, 56182323, 60432113, 22279000, 22279002
1538	87602858 (3071, 3072)	Novel Protein sim. GBank gi 108024 pir E32891 - finger protein 2, placental - human	Contains protein domain (PF00096) - Zinc finger, C2H2 type		transcriptiador	284886, 18108357, 18108394, 21806767, 21806788, 28146628, 35898286, 265020, 265021, 52844150, 264893, 68714117, 29331825, 29331828, 284508, 264905, 20281148, 264808, 18108374, 35898423, 35895855, 265008, 264634, 264838, 264838, 18108385, 58526488, 285017, 285018, 284563, 284782, 18108351, 264448, 264369, 284786
1537	96354558 (3073, 3074)	Novel Protein sim. GBank gi 3876332 emb CA802088 (Z79754) cDNA EST EMBL:T01054 comes from this gene; cDNA EST EMBL:D73600 comes from this gene; cDNA EST yk426112.5 comes from this gene; cDNA EST yk342110.5 comes from this gene; cDNA EST yk475c5.5 comes from this gene; cDNA ES...				65274572, 58182575, 60432049, 284259, 29331828, 285006, 285007, 60433356, 60433438, 264601, 18108351, 284448, 284389, 264288, 33657023, 65274620, 33657109, 80432113
1538	85724828 (3075, 3078)	Novel Protein sim. GBank gi 403440 (M81787) - Gallus domesticus skeletal muscle mRNA, partial cds, l. gene product [Gallus gallus]	Contains protein domain (PF00069) - Eukaryotic protein kinase domain		kinase	18108394, 18108397, 264909, 265008, 285009, 265010, 18108351, 264838, 18108392, 18108385, 18108388

1539	95337628 (3077, 3078)	Novel Protein sim. GBank gij3218411[embjCAA19575.1] - (AL023859) SPBC19C7.07c, putative RNA splicing endonuclease ga mma subunit, len:284aa, similar eg. to YAR008W, YAH8_YEAST, P39707, YAR008W, tRNA splicing endonuclease gamma subunit, (275aa), fasta scores, opt:268, E(1):6.4e-2...		nuclease	22278994, 22278998, 35896288, 58994075, 22278997, 22278998, 22278999, 80432049, 264259, 29331822, 29331824, 60424289, 29331825, 60432289, 29331826, 29331827, 29331828, 35698052, 33858970, 56182435, 265009, 33657402, 60433356, 60433438, 55812038, 21906754, 87188559, 285017, 285018, 265019, 18108351, 264288, 52844229, 18108358, 21906764, 21906787, 21908768, 35695917, 285020, 285021, 52844150, 33857023, 33857109, 27486281, 18108370, 18108378, 35696423, 55811576, 65274791, 284558, 58182323, 80170394, 83373044, 87188518, 60432113, 22278000, 22279002, 264568
1540	95352858 (3079, 3080)	Novel Protein sim. GBank gij5052834[gbjAAD38647.1]AF14567 - (AF145672) BcDNA GH12174 [Drosophila melanogaster]		UNCLASSIFIED	264369, 264691, 263978
1541	95317848 (3081, 3082)	Novel Protein sim. GBank gij5052349[gbjAAD38515.1]AF13501 - (AF135016) protein phosphatase 2A 48 kDa regulatory subunit [Homo sapiens]		phosphatase	264488, 264489, 22278999, 264259, 29331822, 35696052, 264508, 264509, 284905, 284907, 284908, 264511, 264512, 284910, 264592, 264781, 264762, 264448, 264764, 264288, 264687, 21808769, 55811957, 35895917, 265020, 264691, 33657023, 264692, 33857109, 264828, 18108374, 264632, 264634, 264635, 264839, 18108385, 264563, 264564, 264565, 264566, 264486
1542	90937549 (3083, 3084)	Novel Protein sim. GBank gij5305702[gbjAAD41779.1]AF12886 - (AF128887) calpain-like protease [Mus musculus]		calhepsin	18108392, 18108394, 65274572, 29331822, 264508, 265007, 285008, 285009, 265011, 264682, 18108354, 18108355, 52844150, 18108368, 264636, 16108381, 18108382
1543	84348768 (3085, 3086)	Novel Protein sim. GBank gij728832[spjP39189]ALU2_HUMAN - IIII ALU SUBFAMILY SB WARNING ENTRY IIII		nuclease	264106, 33109854, 265019, 264683, 35895917, 284890, 264692, 33657109
1544	87757295 (3087, 3088)	Novel Protein sim. GBank gij3493162 (AF084259) - bromodomain-containing protein BP75 [Mus musculus]	Contains protein domain (PF00439) - Bromodomain	nud_rept	35698286, 56994075, 22278999, 35898052, 60433356, 60433438, 265011, 264663, 33657109, 35698423, 264631, 87186518, 22279000
1545	85757973 (3089, 3090)	Novel Protein sim. GBank gij1088591 (U41007) - similar to S. cerevisiae nuclear protein SNF2 (SP-P22082) in a region of gly-arg repeats [Caenorhabditis elegans]		UNCLASSIFIED	264112, 264692, 264693, 55811576
1546	79478589 (3091, 3092)			UNCLASSIFIED	284905, 264688
1547	86995594 (3093, 3094)	Novel Protein sim. GBank gij2861132 (AF035683) - p21 [Mus musculus]		UNCLASSIFIED	55811366, 285010, 264800, 285017, 285019, 264288, 264768, 265020, 285022, 55811576, 18108380, 264583

1548	94233065 (3095, 3098)	Novel Protein sim. GBank gij3043692[dbj BAA25510] - (AB011156) KIAA0584 protein [Homo sapiens]		UNCLASSIFIED	29331624, 60431526, 264638, 56182323
1549	95330048 (3097, 3098)	Novel Protein sim. GBank gij5689519[dbj BAA83043.1] - (AB029014) KIAA1091 protein [Homo sapiens]		eph	60424178, 22276995, 35696286, 22278998, 22278996, 264092, 264094, 29331622, 58182161, 29331824, 35696052, 264905, 264906, 264908, 264909, 265006, 264511, 265006, 60431735, 60433358, 21906754, 55611386, 87168559, 265017, 265018, 265019, 55811150, 264682, 264286, 264369, 58181562, 264768, 21906765, 21906768, 21906769, 55811957, 265020, 264691, 33657109, 60431528, 35696423, 35695855, 56526486, 60432113, 22279002, 264563, 264566
1550	95201907 (3098, 3100)	Novel Protein sim. GBank gij544483[sp P35350 GUSB_BOVIN - POSSIBLE GUSTATORY RECEPTOR TYPE B (PPRI PROTEIN)]	Contains protein domain (PF00001) - 7 transmembrane receptor (rhodopsin family)	lm7	65274572, 60432289, 265008, 264910, 265011, 265017, 265019, 264768, 56182323
1551	88077111 (3101, 3102)	Novel Protein sim. GBank gij4756566[ref NP_004798.1 pHS6S - heparan-sulfate 6-sulfotransferase		UNCLASSIFIED	22278999, 29331622, 264508, 264509, 264906, 264907, 264908, 265007, 264512, 264910, 21906754, 265018, 265019, 264681, 264764, 264766, 264688, 264769, 21906769, 264692, 35695763, 264635, 264555, 264558, 264557, 264638, 264558, 264563
1552	87617114 (3103, 3104)			UNCLASSIFIED	284259, 29331828, 66712502, 264764, 264286, 264666, 33657109, 264558
1553	94725512 (3105, 3106)	Novel Protein sim. GBank gij4588570[dbj BAA76807.1] - (AB023180) KIAA0983 protein [Homo sapiens]	Contains protein domain (PF00304) - Gamma-thionins family	dehydrogenase	56182575, 35696286, 29146499, 264509, 264907, 264908, 264909, 56182435, 265006, 265008, 265009, 264910, 264757, 264758, 265017, 5811150, 18106351, 264764, 56181562, 35695917, 264693, 33657109, 18108374, 35696423, 65274791, 35695855, 264635, 264555, 56182323, 16108382, 63373044, 22279000
1554	94233069 (3107, 3108)	Novel Protein sim. GBank gij3043692[dbj BAA25510] - (AB011158) KIAA0584 protein [Homo sapiens]	Contains protein domain (PF00446) - Gonadotroph-releasing hormones	transferase	35696286, 22278997, 264259, 29331822, 29331824, 29331625, 29331828, 265007, 265009, 60432229, 33657402, 55812038, 265011, 265019, 264681, 264369, 264686, 264767, 264768, 21906765, 21906769, 35695917, 264693, 18108370, 60431526, 55611576, 264631, 60170394, 56182323, 63373044, 16108385, 22279000, 22279002

1555	8732970 (3109, 3110)	Novel Protein sim. GBank gij2257485[dbj]BAA213921 - (AB004534) p1015 [Schizosaccharomyces pombe]	Contains protein domain (PF00400) - UNCLASSIFIED WD domain, G-beta repeat	264259, 29331826, 35696052, 264508, 264905, 264908, 264907, 264906, 52844045, 264909, 264910, 80432228, 60433356, 55812038, 264758, 264759, 33657084, 265011, 87168559, 264801, 265018, 265019, 264763, 264764, 264288, 264768, 264768, 21908765, 35895917, 265022, 264691, 33657023, 35898423, 35895855, 264635, 264555, 264838, 264638, 264639, 16108365, 56528488
1556	91228268 (3111, 3112)		UNCLASSIFIED	83373044, 264758, 265022, 264600, 35696052, 264630, 35696423, 265018, 264632, 264682, 29331822, 265020, 265011
1557	87640809 (3113, 3114)	Novel Protein sim. GBank gij3328611 (AF078783) - contains similarity to C3HC4-type zinc fingers (Pfam; z1-C3HC4.hmm, score: 34.08); most similar to drosophila goliath protein (SW: Q06003) [Caenorhabditis elegans]	Contains protein domain (PF000097) - UNCLASSIFIED Zinc finger, C3HC4 type (RING finger)	80432228, 264509, 264806, 264907, 264908, 264909, 264910, 264758, 55811386, 264761, 264762, 264768, 264769, 264890, 263978, 264834, 264835, 264639, 264584, 264488
1558	94840376 (3115, 3116)	Novel Protein sim. GBank gij5360105[gb]A042671.1[AF155105] putative zinc finger protein NY-REN-34 antigen [Homo sapiens]		22278994, 22278998, 22278987, 22278998, 22278999, 60432048, 264259, 29331824, 29331825, 29331826, 29331827, 264908, 264909, 60433356, 21908754, 265017, 265018, 264448, 21908767, 265021, 265022, 33657023, 33657109, 18108370, 55811578, 83373044, 87188518, 22278900, 22279002
1559	86224865 (3117, 3118)	Novel Protein sim. GBank gij112908[sp]P02750[A2GL_HUMAN - LEUCINE-RICH ALPHA-2-GLYCOPROTEIN (LRG)]	Contains protein domain (PF00560) - glycoprotein Leucine Rich Repeat	264259, 29331822, 60432289, 35896052, 264107, 264110, 21906754, 33108954, 87188559, 264780, 264763, 21906764, 21908765, 21908769, 265021, 264890, 35895855
1560	84580875 (3119, 3120)	Novel Protein sim. GBank gij3880146[emb]CAA827041 - (Z88319) Similarity to Human hnRNP F protein (PIR Acc. No. S43484); cDNA EST EMBL:D34218 comes from this gene; cDNA EST EMBL:D37248 comes from this gene; cDNA EST EMBL:D71617 comes from this gene; cDNA EST EMBL:D74531 comes from...	UNCLASSIFIED	264908, 264603, 264638
1561	86608159 (3121, 3122)			
1562	83358682 (3123, 3124)		UNCLASSIFIED	264510, 264594
1563	85508694 (3125, 3126)		UNCLASSIFIED	263967
1564	87766371 (3127, 3128)	Novel Protein sim. GBank gij1188287[sp]P45953[JACOV_RAT - ACYL-COA DEHYDROGENASE, VERY-LONG-CHAIN SPECIFIC PRECURSOR (VLCAD)]	Contains protein domain (PF00441) - dehydrogenase Acyl-CoA dehydrogenase	16108394, 35696288, 264259, 29331822, 80432289, 35898052, 29331828, 264508, 58712502, 264908, 58182435, 265007, 264910, 80170631, 21908754, 265011, 265017, 265018, 265019, 18108351, 264448, 264288, 264686, 21908765, 21908768, 265021, 80170815, 264892, 35898423, 35895855, 264557, 58182323, 80432113, 22278902, 264482

1565	87783381 (3128, 3130)	Novel Protein sim. GBank g 129726 p P05307 PDI_BOVIN - PROTEIN DISULFIDE ISOMERASE PRECURSOR (PDI) (PROLYL 4-HYDROXYLASE BETA SUBUNIT) (CELLULAR THYROID HORMONE BINDING PROTEIN) (P55)		isomerase	264488, 264888, 18108398, 55811857, 264534, 264259, 284508, 264905, 264509, 264906, 18108372, 284510, 284511, 284512, 265008, 264630, 265009, 284910, 264635, 264638, 264591, 264555, 264592, 264637, 264593, 264594, 264595, 264596, 265011, 264603, 22279002, 18108351, 264762, 264565, 264567
1566	87424749 (3131, 3132)	Novel Protein sim. GBank g 3880445 emb CAA20329 - (ALD31266) VM106R.1 [Caenorhabditis elegans]		Inf	22278998, 22278998, 264259, 29331822, 29331824, 60432289, 29331827, 66712502, 264908, 265006, 18108351, 52644229, 21908765, 21908787, 21908769, 21908769, 33657109, 264555, 264639, 264462
1567	84999008 (3133, 3134)	Novel Protein sim. GBank g 4929699 gb AAD34110.1 AF15167 - (AF151873) CGI-115 protein [Homo sapiens]		UNCLASSIFIED	56162575, 21908769, 264692
1568	87648781 (3135, 3136)	Novel Protein sim. GBank g 4827063 ref NP_005072.1 pZNF1 - zinc finger protein 142 (clone pTZ-49)	Contains protein domain (PF000096) - Zinc finger, C2H2 type	transcriptfactor	29331827, 29331830, 264511, 265009, 284758, 21906767, 21906788, 264691, 264693, 22279000, 22279002
1569	90936668 (3137, 3138)	Novel Protein sim. GBank g 5689451 p BAA83009.1 - (AB028980) KIAA1057 protein [Homo sapiens]	Contains protein domain (PF00443) - Ubiquitin carboxyl-terminal hydrolase family 2	ubiquitin	85274572, 29331822, 29331824, 29331828, 264905, 56182435, 265007, 265019, 264764, 21908765, 21908788, 55811857, 60170615, 52644150, 284892, 33657023, 33657109, 18108377, 264563, 264567
1570	88943981 (3139, 3140)	Novel Protein sim. GBank g 1255430 U53155) - No definition line found [Caenorhabditis elegans]		UNCLASSIFIED	264595, 264682, 265021
1571	91210340 (3141, 3142)	Novel Protein sim. GBank g 4507731 ref NP_001061.1 pTUBG - tubulin, gamma polypeptide	Contains protein domain (PF000081) - Tubulin/FilZ family	tubulin	22278996, 35696286, 22278997, 264091, 264259, 29331824, 29331825, 29331827, 35698052, 284508, 264905, 56182435, 264510, 265007, 264756, 265011, 18108351, 264448, 264288, 264369, 21908765, 21908767, 21908788, 21908789, 35695917, 265020, 265021, 33657023, 264693, 18108370, 18108377, 35696423, 35695855, 264634, 264555, 284558, 18108384
1572	81201664 (3143, 3144)			UNCLASSIFIED	35696052, 264905, 264908, 264908, 264910, 264758, 264766, 35695917, 264637
1573	80207066 (3145, 3146)			UNCLASSIFIED	263972

1574	94216142 (3147, 3148)	Novel Protein sim. GBank gi14758334 refNP_004256.1 pFADS - delta-6 fatty acid desaturase	Contains protein domain (PF00173) - Heme-binding domain in cytochrome b5 and oxidoreductases	cytochrome	18108394, 264887, 18108397, 18108398, 22278996, 22278997, 22278999, 264259, 29331825, 29331827, 29146498, 29146499, 264107, 264907, 264909, 52644045, 264511, 285008, 264910, 285009, 284591, 21908754, 285011, 265019, 18108351, 284882, 284763, 264764, 18108354, 264389, 264288, 264885, 284768, 264688, 264768, 264688, 21906765, 21908788, 21908787, 21906788, 21906789, 29148629, 264690, 264691, 284693, 20281069, 18108370, 18108374, 18108379, 35895855, 264834, 18108384, 18108385, 22278902, 264563, 284568
1575	95340019 (3148, 3150)	Novel Protein sim. GBank gi1388181 emb CAA94858 - (Z70783) similar to EF-hand calcium binding protein; cDNA EST EMBL: C08700 comes from this gene [Caenorhabditis elegans]	Contains protein domain (PF00036) - EF hand	phosphatase	56994075, 284258, 28331822, 28331824, 29331825, 80432289, 29331828, 264909, 265006, 265008, 285009, 265010, 87168559, 55811150, 264448, 18108354, 284369, 264288, 18108357, 55811957, 265020, 265021, 60170615, 264691, 33657023, 33657109, 60431528, 85274791, 35695855, 18108385, 60432113, 22279002, 264482
1576	95314019 (3151, 3152)	Novel Protein sim. GBank gi12773195 (AF038711) - contains similarity to Physcomitrella patens glyceraldehyde 3-phosphate dehydrogenase (GB:X72381) [Caenorhabditis elegans]	UNCLASSIFIED	UNCLASSIFIED	264569, 264092, 264094, 264095, 264259, 264508, 264905, 264509, 284907, 264909, 284510, 264511, 265008, 284910, 21908754, 265010, 265011, 87168559, 264761, 284762, 264288, 264768, 264789, 264891, 264893, 35895855, 264632, 264834, 264635, 264838, 63373044, 264486
1577	87613800 (3153, 3154)	Novel Protein sim. GBank gi12499130 sp P70315 WASP_MOUSE - WISKOTT - ALDRICH SYNDROME PROTEIN HOMOLOG (WASP)	UNCLASSIFIED	UNCLASSIFIED	264511, 265011, 264681, 284369, 264688, 264689, 264829, 264555, 264558, 264559
1578	87123138 (3155, 3156)				264259, 29331826, 265017, 264689, 264893, 60432113
1579	88085141 (3157, 3158)	Novel Protein sim. GBank gi12978255 dbj BAA25190 - (AB007407) myeloid zinc finger protein-2 [Mus musculus]	Contains protein domain (PF00086) - Zinc finger, C2H2 type	UNCLASSIFIED	35686288, 264908, 264909, 60433438, 55811388, 264369, 264685, 33857023, 284555, 284558, 284557, 87168518
1580	87255702 (3159, 3160)	Novel Protein sim. GBank gi14324682 bbj AAD16986 - (AF108674) late gestation lung protein 1 [Rattus norvegicus]	Contains protein domain (PF00188) - SCP-like extracellular protein	glycoprotein	22278999, 35698052, 29331830, 52644045, 55812038, 87168474, 285018, 204448, 265022, 264838, 56526488, 22279000
1581	95087431 (3161, 3182)	Novel Protein sim. GBank gi12088838 (AF003386) - F59E12.4 gene product [Caenorhabditis elegans]			22278995, 29331822, 29331824, 29331828, 56182435, 284595, 55812038, 87168559, 285017, 264288, 21908784, 55811957, 35895917, 264692, 55811578, 264637, 56182323, 264559, 83373044, 80432113

1582	85358052 (3163, 3164)	Novel Protein sim. GBank gij5420387[emb]CAB46678.1] - (AJ243459) proteophosphoglycan [Leishmania major]		phosphatase	284258, 60432289, 29331827, 284509, 284905, 284906, 284907, 284909, 284910, 284782, 284288, 284786, 284789, 284632, 284555, 284839, 58528486, 22279000
1583	87622715 (3165, 3186)	Novel Protein sim. GBank gij5578959[emb]CAB51351.1] - (AL050306) d.47587.2 (novel protein) [Homo sapiens]		UNCLASSIFIED	60170831, 33857402, 284882, 21806786, 35695855, 284563
1584	85337722 (3167, 3168)	Novel Protein sim. GBank gij5531615[gb]AAD44482.1] - (AF078850) steroid dehydrogenase homolog [Homo sapiens]	Contains protein domain (PF00106) - short chain dehydrogenase	dehydrogenase	60424179, 52848842, 65274572, 56182575, 22278995, 35686286, 22278998, 22278998, 22278999, 284258, 29331822, 56182181, 60424289, 60432289, 29331827, 29331828, 35696052, 28146488, 66712502, 29331830, 52844045, 58182435, 284510, 284512, 265008, 60433358, 33857402, 80433438, 55812038, 21906754, 55811386, 52844298, 87188474, 87188559, 285018, 265019, 264448, 264389, 284288, 18108358, 21908765, 21906767, 21908788, 21906768, 35695917, 285020, 285021, 285022, 52844150, 33657023, 33657108, 18108374, 55810784, 55811578, 35696423, 65274781, 35695855, 56182323, 83373044, 18108387, 87188518, 60432113, 22279002
1585	87626117 (3168, 3170)	Novel Protein sim. GBank gij4240132[dbj]BAA74848.1] - (AB020630) KIAA0823 protein [Homo sapiens]	Contains protein domain (PF00023) - Ank repeat	phosphatase	35696288, 22278998, 284259, 29331822, 29331824, 29331825, 264805, 285008, 265007, 265008, 60433356, 33109954, 87188474, 285011, 265017, 284804, 264369, 264288, 284885, 284769, 18108359, 21906765, 18108384, 18108370, 284828, 263972, 18108383, 18108388, 264482, 284584
1586	88067081 (3171, 3172)	Novel Protein sim. GBank gij3788484 (AF098993) - No definition line found [Caenorhabditis elegans]			265017, 265018, 264689, 33657023, 263978, 284636, 264563
1587	87617128 (3173, 3174)	Novel Protein sim. GBank gij3253159 (AF003355) - translation initiation factor eIF2C [Oryctolagus cuniculus]		UNCLASSIFIED	284907, 284908, 264511, 284910, 284591, 264594, 284828, 264831, 264583, 264483, 284587
1588	87802536 (3175, 3176)	Novel Protein sim. GBank gij1077573[pir]S52680 - probable ribosomal protein L34, mitochondrial - yeast (Saccharomyces cerevisiae)	Contains protein domain (PF00468) - Ribosomal protein L34	UNCLASSIFIED	264259, 29331828, 284905, 265006, 264758, 21906754, 284781, 284782, 21906785, 21908788, 60170815, 52844150, 33657109, 35695855, 58182323, 18108385
1589	90980883 (3177, 3178)	Novel Protein sim. GBank gij2137756[pir]J48746 - semaphorin C - mouse (fragment)		UNCLASSIFIED	65274572, 284490, 29331822, 68714117, 29331827, 29331828, 58182435, 265008, 60170831, 284595, 264758, 284598, 265011, 284688, 21908766, 21906768, 55811857, 27488265, 284839, 18108385, 56526488, 80432113

1590	95318825 (3179, 3180)			UNCLASSIFIED	264489, 22278996, 264259, 29331824, 29331825, 29331828, 29331827, 265006, 60433356, 21906754, 285017, 265018, 265019, 264448, 284785, 264288, 52844228, 21906765, 21906767, 21906788, 21906789, 285021, 264892, 27468265, 35695763, 55528488, 60432113, 22279000, 22279002, 264564
1591	86877160 (3181, 3182)				
1592	87882533 (3183, 3184)			MHC	264259, 264905, 29331830, 264595, 265017, 264448, 264288, 264690, 264629, 87188518
1593	94891661 (3185, 3186)		Novel Protein sim. GBank gi 4557749 ref NP_000237.1 pMHC2 - MHC class II transactivator		
1594	87773752 (3187, 3188)		Novel Protein sim. GBank gi 3877072 emb CAA87060 - (Z46937) similarity with ribosomal protein L21 (Caenorhabditis elegans)	UNCLASSIFIED	65274572, 60432048, 264509, 60433358, 21908754, 21908767, 21908768, 16108370, 35698423, 22279000, 264555, 264587
1595	78919425 (3189, 3190)		Novel Protein sim. GBank gi 3152703 (AF085389) - tetraspan NET-4 [Homo sapiens]	UNCLASSIFIED	264486, 29331827, 264905, 264906, 264907, 264908, 264909, 264910, 264592, 264593, 284757, 264602, 264604, 264760, 264681, 264288, 264768, 264768, 29146629, 35695917, 264692, 264628, 264629, 264630, 264632, 264834, 264635, 264636, 264639, 264563, 264564, 264566
1596	78933928 (3191, 3192)			Contains protein domain (PF00335) - 4 transmembrane segments integral membrane proteins	29331826, 264908, 55811957
1597	88871857 (3193, 3194)		Novel Protein sim. GBank gi 5257114 gb AAD41244.1 AF094480 cholesterol 24-hydroxylase [Homo sapiens]	UNCLASSIFIED cyto450	29146498, 264756, 283967 264092, 29331824, 264508, 264682, 264369, 264688, 264830, 264583
1598	87862839 (3195, 3196)				264259, 284634
1599	87848629 (3197, 3198)		Novel Protein sim. GBank gi 4506797 ref NP_000324.1 pSCA7 - spinocerebellar ataxia 7 (olivopontocerebellar atrophy with retinal degeneration)	UNCLASSIFIED	52845080, 29331824, 29331826, 264511, 265009, 265011, 264605, 264446, 264764, 285020, 264692, 264693, 18108370, 264835, 18108385
1600	80056002 (3199, 3200)				
1601	15023246 (3201, 3202)			UNCLASSIFIED	29331826, 264603, 264691, 264563
1602	86926987 (3203, 3204)		Novel Protein sim. GBank gi 5305704 gb AAD41780.1 AF128535 cytoplasmic phosphoprotein PACSIN2 [Mus musculus]	264635	
1603	80502072 (3205, 3206)		Novel Protein sim. GBank gi 263920 ref S27839 - tenascin - chicken	UNCLASSIFIED struct	29146499, 264112, 264762, 18108351, 29146827, 263974
1604	80221813 (3207, 3208)		Novel Protein sim. GBank gi 4786831 gb AAD28633.1 AF116827 - unknown [Homo sapiens]	collagen	264490, 29331824, 264807, 264909, 264511, 265008, 264592, 265010, 285011, 264782, 264784, 284389, 264288, 264887, 264789, 264693, 264626, 264834, 264638, 264555, 264556, 264838, 264557, 264558, 264559, 18108365
				ATPase-associated	263977

1605	91221129 (3208, 3210)				264905, 264509, 264908, 264907, 264908, 264909, 264604, 264766, 264768, 264682, 264683, 33657109, 264629, 35695855, 264635, 264636, 264637	struct		
1606	94312703 (3211, 3212)	Novel Protein sim. GBank gij4505313[ref]NP_003794.1pMYOM - UNKNOWN	Contains protein domain (PF000047) - Immunoglobulin domain		22278998, 22278998, 264259, 33657402, 265017, 18108351, 264446, 21908767, 21906769, 52844150, 264691, 87186516, 264689	struct		
1607	10871805 (3213, 3214)	Novel Protein sim. GBank gij5174473[ref]NP_005886.1[pp] - Intracisternal A particle-promoted polypeptide				transcriptfactor		
1608	80428900 (3215, 3216)	Novel Protein sim. GBank gij2224629[db][BA20802] - (AB002342) KIAA0344 [Homo sapiens]				UNCLASSIFIED		
1609	94311572 (3217, 3218)	Novel Protein sim. GBank gij4884073[emb][CA843213.1] - (ALC049834) hypothetical protein [Homo sapiens]			52844507, 52845156, 52846385, 52846842, 56182575, 22278994, 56994075, 35696288, 22278997, 22278998, 22278999, 264259, 52845080, 29147620, 29331628, 35698052, 33656970, 264508, 264509, 264907, 52844045, 56182435, 284510, 284511, 264512, 33657402, 21906754, 52846317, 33108954, 52844296, 87168474, 285017, 265018, 265019, 18108351, 264446, 264286, 284769, 52844229, 21906765, 21906766, 21906767, 21906768, 21906768, 55611857, 35695917, 265020, 265021, 265022, 52844150, 33657023, 33657109, 52845129, 27486261, 27486262, 35695763, 264626, 18108370, 18108376, 35698423, 264638, 52844332, 18108387, 87188518, 22279000, 264583, 264466	UNCLASSIFIED		
1610	85466200 (3219, 3220)	Novel Protein sim. GBank gij263920[pr][S27839 - lensin - chicken				UNCLASSIFIED		
1611	94122843 (3221, 3222)	Novel Protein sim. GBank gij107284[pr][A35415 - peroxidase (EC 1.11.1.7), thyroid (Grave's disease) - human (fragment)	Contains protein domain (PF00006) - peroxidase EGF-like domain		264593, 264757, 55612038, 265018, 265020, 264691, 264692, 264693, 264631, 264634, 264635, 264555, 22279000, 264564, 35696286, 21906765, 264691, 35696423	peroxidase		
1612	65746031 (3223, 3224)	Novel Protein sim. GBank gij387484[emb][CAA94337] - (Z70307) Similarity to B.subtilis tetracycline resistance protein (SW:TCP2_BACSU); cDNA EST EMBL:CO8951 comes from this gene; cDNA EST EMBL:CO8285 comes from this gene [Caenorhabditis elegans]				UNCLASSIFIED		
1613	62247354 (3225, 3226)					UNCLASSIFIED		264759

1614	81228634 (3227, 3228)	Novel Protein sim. GBank gi4680673 gb AAD27726.1 AF13295 - (AF13295.1) CGI-17 protein [Homo sapiens]	Contains protein domain (PF01605) - ERF 1-like proteins	UNCLASSIFIED	22278995, 22278996, 22278997, 22278998, 22278999, 264259, 29331822, 264906, 264512, 265009, 265011, 265017, 265018, 265019, 16108351, 264683, 264288, 264766, 21906767, 21908768, 21906789, 35695917, 265021, 265022, 35696423, 35695655, 60170394, 58182323, 83373044, 264558
1615	86121909 (3228, 3230)	Novel Protein sim. GBank gi5688485 db BAAB3026.1 - (AB028897) KIAA1074 protein [Homo sapiens]	Contains protein domain (PF00023) - Ank repeat	homeobox	22278996, 35696288, 22278997, 29331822, 35696052, 29331828, 284508, 264908, 284909, 56182435, 264511, 265017, 265018, 264768, 264767, 264788, 265020, 264891, 264628, 264632, 264635, 264555, 264558, 56182323, 284558, 22279002
1616	94311819 (3231, 3232)	Novel Protein sim. GBank gi3878260 emb CAB01696 - (Z78416) cDNA EST EMBL:D71020 comes from this gene; cDNA EST EMBL:D73593 comes from this gene; cDNA EST EMBL:C07849 comes from this gene; cDNA EST EMBL:C09081 comes from this gene; cDNA EST YK39972.3 comes from this gene; cDNA ...	UNCLASSIFIED	UNCLASSIFIED	264468, 52844507, 52845158, 52846365, 52846842, 22278994, 22278995, 35698286, 22278996, 22278997, 22278999, 52645080, 29331822, 29331824, 29331825, 29331827, 29331828, 35696052, 33856970, 264905, 264909, 264594, 52848317, 21906754, 33857064, 52844296, 87188474, 67188559, 265017, 265018, 265019, 264681, 264448, 264684, 52844229, 21906764, 264889, 21906765, 21906768, 21906769, 35695917, 265020, 265021, 52844150, 33857023, 27486262, 33657349, 27486285, 35695783, 52845129, 33657109, 33657182, 27486281, 18108378, 35696423, 35695855, 284557, 52844332, 264558, 18108385, 67168518
1617	88090742 (3233, 3234)	Novel Protein sim. GBank gi466053 sp P34679 YO41_CAEEL - HYPOTHETICAL 66.7 KD PROTEIN ZK757.1 IN CHROMOSOME III	Contains protein domain (PF01529) - DHC zinc finger domain	peptidase	35696052, 264905, 264509, 264907, 264908, 264510, 284511, 264784, 264786, 264788, 264689, 264693, 18106374, 264635, 264638, 264638
1618	86272860 (3235, 3236)	Novel Protein sim. GBank gi424023 db BAAT4894.1 - (AB020676) KIAA0871 protein [Homo sapiens]		struc1	35696286, 22278999, 264092, 29331824, 29331825, 35696052, 33657094, 21906765, 27488284
1619	95354560 (3237, 3238)	Novel Protein sim. GBank gi5031763 ref NP_005515.1 pHRY - hairy (Drosophila)- homolog	Contains protein domain (PF00010) - Helix-loop-helix DNA-binding domain	transcriptfactor	52846842, 65274572, 22278999, 264259, 29331822, 29331824, 29331825, 29331826, 29331827, 29331828, 35696052, 58182435, 265007, 285006, 284910, 60170631, 60432229, 60433358, 60433438, 265019, 264446, 284286, 264686, 21908788, 265021, 60170615, 33657023, 65274620, 33657109, 18108374, 18108378, 35696423, 35695655, 58182323, 58526468
1620	87344655 (3239, 3240)	Novel Protein sim. GBank gi1351047 sp P45843 SCRT_DROME - SCARLET PROTEIN		UNCLASSIFIED	264684

1821	87078708 (3241, 3242)	Novel Protein sim. GBank g 3822553 (AF098788) - nuclear calmodulin-binding protein [Gallus gallus]	Contains protein domain (PF00822) - SPRY domain	UNCLASSIFIED	264910	18108392, 65274572, 18108398, 22278996, 22278997, 22278998, 29146498, 29146499, 264905, 264908, 264909, 264928, 52644045, 264592, 80433358, 21906754, 284602, 265017, 264368, 21908768, 55811957, 265021, 80170615, 264635, 264557, 60170394, 83373044, 18108385, 22279000, 22279002, 284566
1822	84741738 (3243, 3244)	Novel Protein sim. GBank g 731088 p 40388 UV22_SCHPO - UV-INDUCED PROTEIN UV22		UNCLASSIFIED	264910	18108392, 65274572, 18108398, 22278996, 22278997, 22278998, 29146498, 29146499, 264905, 264908, 264909, 264928, 52644045, 264592, 80433358, 21906754, 284602, 265017, 264368, 21908768, 55811957, 265021, 80170615, 264635, 264557, 60170394, 83373044, 18108385, 22279000, 22279002, 284566
1823	87779108 (3245, 3246)	Novel Protein sim. GBank g 731088 p 40388 UV22_SCHPO - UV-INDUCED PROTEIN UV22		ribosomalprot	264910	18108392, 65274572, 18108398, 22278996, 22278997, 22278998, 29146498, 29146499, 264905, 264908, 264909, 264928, 52644045, 264592, 80433358, 21906754, 284602, 265017, 264368, 21908768, 55811957, 265021, 80170615, 264635, 264557, 60170394, 83373044, 18108385, 22279000, 22279002, 284566
1824	87338178 (3247, 3248)	Novel Protein sim. GBank g 3875668 emb CAB05478 - (283104) cDNA EST EMBL: T00015 comes from this gene; cDNA EST EMBL: D33665 comes from this gene; cDNA EST EMBL: D36540 comes from this gene; cDNA EST yk2408.3 comes from this gene; cDNA EST yk387c8.3 comes from this gene; cDNA ES...		UNCLASSIFIED	264910	18108392, 65274572, 18108398, 22278996, 22278997, 22278998, 29146498, 29146499, 264905, 264908, 264909, 264928, 52644045, 264592, 80433358, 21906754, 284602, 265017, 264368, 21908768, 55811957, 265021, 80170615, 264635, 264557, 60170394, 83373044, 18108385, 22279000, 22279002, 284566
1825	95354748 (3249, 3250)	Novel Protein sim. GBank g 4589622 cb BAA76833.1 - (AB023206) KIAA0989 protein [Homo sapiens]		kinase	264910	18108392, 65274572, 18108398, 22278996, 22278997, 22278998, 29146498, 29146499, 264905, 264908, 264909, 264928, 52644045, 264592, 80433358, 21906754, 284602, 265017, 264368, 21908768, 55811957, 265021, 80170615, 264635, 264557, 60170394, 83373044, 18108385, 22279000, 22279002, 284566
1826	94754369 (3251, 3252)	Novel Protein sim. GBank g 5878070 gb AAD4844.1 AF16090 - (AF160904) BcDNA HL05938 [Drosophila melanogaster]			264910	18108392, 65274572, 18108398, 22278996, 22278997, 22278998, 29146498, 29146499, 264905, 264908, 264909, 264928, 52644045, 264592, 80433358, 21906754, 284602, 265017, 264368, 21908768, 55811957, 265021, 80170615, 264635, 264557, 60170394, 83373044, 18108385, 22279000, 22279002, 284566
1827	83368773 (3253, 3254)	Novel Protein sim. GBank g 3668087 (AC004667) - hypothetical protein [Arabidopsis thaliana]		UNCLASSIFIED	264910	18108392, 65274572, 18108398, 22278996, 22278997, 22278998, 29146498, 29146499, 264905, 264908, 264909, 264928, 52644045, 264592, 80433358, 21906754, 284602, 265017, 264368, 21908768, 55811957, 265021, 80170615, 264635, 264557, 60170394, 83373044, 18108385, 22279000, 22279002, 284566
1828	85708459 (3255, 3256)	Novel Protein sim. GBank g 3668087 (AC004667) - hypothetical protein [Arabidopsis thaliana]		eph	264910	18108392, 65274572, 18108398, 22278996, 22278997, 22278998, 29146498, 29146499, 264905, 264908, 264909, 264928, 52644045, 264592, 80433358, 21906754, 284602, 265017, 264368, 21908768, 55811957, 265021, 80170615, 264635, 264557, 60170394, 83373044, 18108385, 22279000, 22279002, 284566

1629	84993841 (3257, 3258)	Novel Protein sim. GBank gij4240175jdbj[BAA74866.1] - (AB020850) KIAA0843 protein [Homo sapiens]		struct	264555
1630	87778027 (3259, 3260)			UNCLASSIFIED	29331822, 29331827, 285010, 264683, 284634, 22279002
1631	87758454 (3261, 3262)	Novel Protein sim. GBank gij1915692jemb[CAA69995] - (Y08740) tom-1A protein [Gallus gallus]		UNCLASSIFIED	55811857, 284258, 33657023, 264893, 29331822, 29331824, 29331827, 29331828, 284906, 264808, 55811578, 284910, 264834, 284636, 284637, 56182323, 284559, 264758, 18108385, 284563, 284764, 284766
1632	87871692 (3263, 3264)	Novel Protein sim. GBank gij2558501jdbj[BAA22898] - (D63850) hepatoma-derived growth factor [Mus musculus]		UNCLASSIFIED	264687, 284788, 284891, 284692, 29146499, 284508, 264805, 284907, 284511, 264512, 264482, 264681, 284763, 264882, 264883, 284488, 284259, 284907, 284908, 264809, 284628, 264828, 264631
1633	87773883 (3265, 3266)				265007, 284637, 22278002
1634	85992817 (3267, 3268)	Novel Protein sim. GBank gij4887229jdbj[AAD32244.1]AF15075 - (AF150755) microtubule-actin crosslinking factor [Mus musculus]	Contains protein domain (PF00435) - spectrin repeat	struct	
1635	94232800 (3269, 3270)			UNCLASSIFIED	65274572, 22278996, 35696052, 52644045, 264511, 265008, 285008, 265010, 285011, 265018, 265019, 284448, 284369, 21908785, 21908788, 265021, 284890, 284482
1636	80413227 (3271, 3272)	Novel Protein sim. GBank gij455751jdbj[NP_001339.1]pDAPK - death-associated protein kinase 3	Contains protein domain (PF00069) - Eukaryotic protein kinase domain	UNCLASSIFIED	22278955, 264594, 264763, 265020, 284558
1637	80070435 (3273, 3274)	Novel Protein sim. GBank gij3420051 (AC004880) - unknown protein [Arabidopsis thaliana]			
1638	87101854 (3275, 3276)				21908785, 21908787, 22278988, 35698288, 22278989, 264259, 264682, 264693, 29331824, 33857109, 284508, 284906, 18108370, 264629, 265007, 33857402, 21908754, 284602, 264604, 264764, 264663, 264566, 264288
1639	94322184 (3277, 3278)	Novel Protein sim. GBank gij5420389jemb[CA846680.1] - (A243460) proteophosphoglycan [Leishmania major]		UNCLASSIFIED	264488, 16108394, 65274572, 56182575, 35896286, 29331824, 29331826, 29331827, 35896052, 264906, 56182435, 264112, 265008, 285008, 264757, 264758, 55811386, 264603, 264760, 18108351, 264764, 264288, 264786, 284766, 21908767, 55811957, 264691, 33857023, 85274620, 18108370, 55810784, 55811578, 264558, 264639, 83373044, 18108385, 87188518

1640	94143165 (3279, 3280)	Novel Protein sim. GBank gi 2042469 emb CAA16847.1 - (AL021747) hypothetical protein [Schizosaccharomyces pombe]		UNCLASSIFIED	56182575, 56994075, 35696266, 80432049, 60432289, 29331827, 35696052, 52644045, 56182435, 284510, 285006, 285007, 285008, 284910, 265009, 33657402, 55812038, 285010, 265011, 265017, 265018, 264288, 52644229, 21908765, 21908768, 21908768, 35695917, 265021, 60170615, 52644150, 33857023, 33857109, 33857349, 18108374, 35698423, 65274791, 35695955, 284632, 284555, 56182323, 22279000
1641	87625160 (3281, 3282)			UNCLASSIFIED	29148499, 285006, 285007, 285008, 285009, 55812038, 265010, 285011, 284555, 284558, 284558, 18108383
1642	94312557 (3283, 3284)	Novel Protein sim. GBank gi 1575333 (U60416) - myr 6 myosin heavy chain [Rattus norvegicus]	Contains protein domain (PF01643) - DIL domain	- struct	22278899, 29147820, 29331826, 29331828, 33858970, 55812038, 265010, 265018, 265019, 18108351, 284689, 285020, 285022, 264690, 33657023, 65274620, 35695763, 52644332, 18108381, 60170394, 56182323, 16108388, 67188518, 22279002, 284564
1643	94131768 (3285, 3286)				29331825, 29331827, 29331828, 21908754, 265019, 284288, 284693, 33857349, 18108370, 18108378, 284555, 83373044, 22278002, 284482
1644	88095125 (3287, 3288)			UNCLASSIFIED	284905, 284907, 284908, 284910, 285009, 264757, 264756, 264761, 264782, 264783, 264788, 264786, 264769, 284628, 284829, 284630, 284631, 284632, 284563, 284564, 264565, 284566, 264567
1645	95013858 (3289, 3290)			UNCLASSIFIED	284685, 284693
1648	95382891 (3291, 3292)	Novel Protein sim. GBank gi 1076802 pt J S49915 - extensin like protein - maize		UNCLASSIFIED	22278594, 56994075, 35698288, 264259, 29331824, 29331825, 29331828, 80432289, 284508, 80433358, 80433438, 87168559, 285018, 284687, 35695917, 264692, 33657023, 33657182, 27488281, 27486265, 33657349, 60432113, 264563, 264564
1647	94278428 (3293, 3294)	Novel Protein sim. GBank gi 5002573 emb CAB44338.1 - (Y17468) alpha-N-acetylglucosamine alpha-2,6- sialyltransferase [Fugu rubripes]		UNCLASSIFIED	29331822, 264906, 264908, 264369, 21908788, 60170615, 264639, 22279000
1648	87842098 (3295, 3296)		Contains protein domain (PF00096) - Zinc finger, C2H2 type	UNCLASSIFIED	265009, 284686, 55811957, 35695917, 55810784, 284558, 56182323, 264558, 18108385

1648	85347628 (3287, 3288)	Novel Protein sim. GBank gj1854085[emb]CAA58337] - (XG3413) U88 [Human herpesvirus 6]		cadherin	264488, 22278995, 35696286, 22278996, 22278997, 22278998, 60432049, 264259, 29331822, 29331824, 29331825, 29331826, 29331827, 29331828, 284905, 284907, 68712502, 264908, 52844045, 284909, 56182435, 264511, 265007, 265008, 265009, 264591, 264593, 60433438, 284596, 55812038, 21808754, 285011, 264801, 284802, 285017, 285018, 285019, 284882, 264448, 264764, 264883, 284288, 264768, 264685, 264687, 264788, 264688, 264789, 52844228, 284689, 21808785, 21808788, 21808787, 21808788, 55811957, 35695917, 265021, 265022, 52844150, 284692, 33857023, 33637109, 20281149, 18108370, 284828, 18108374, 18108376, 35696423, 35695855, 284632, 284834, 284635, 284838, 18108380, 284639, 284558, 18108382, 18108384, 18108385, 18108387, 284080, 264404, 60432113, 22278000, 22278002, 264482, 284565, 264566, 264487, 265011, 284802, 21808767, 18108374, 18108377, 18108385
1650	87418539 (3289, 3300)	Novel Protein sim. GBank gj13647335[emb]CAA21059] - (AL031644) possible zinc-finger protein			
1651	91638773 (3301, 3302)	[Schizosaccharomyces pombe] Novel Protein sim. GBank gj14884278[emb]CAB43247.1] - (AL050037) hypothetical protein [Homo sapiens]		synthase	284488, 52845156, 18108397, 35696286, 22278998, 22278999, 264259, 29331822, 29331824, 29331825, 29331826, 29331827, 29331828, 264508, 264908, 29331830, 264910, 60432229, 21808754, 265010, 285011, 285017, 265019, 284448, 18108354, 284288, 264888, 21808765, 21808768, 21808788, 21808769, 265022, 264892, 264693, 264629, 35695855, 264558, 264837, 284557, 284559, 83373044, 56526488, 22279000, 22279002, 284564, 22278997, 28146498, 56182435, 21808754, 264369, 21808765, 21808768, 21808769, 265020, 52844150, 33857109, 22278000, 22278002
1652	86598622 (3303, 3304)	Novel Protein sim. GBank gj1657837 [U73200] - p116Rip [Mus musculus]	Contains protein domain (PF00169) - PH domain	siucl	18108398, 22278995, 22278998, 264259, 29331822, 29331824, 60432289, 29331828, 29331827, 29331830, 284909, 265006, 285009, 80432229, 80433356, 60433438, 21808754, 265017, 285019, 264448, 284683, 264288, 285021, 285022, 264892, 18108384, 60432113, 284587, 33857109, 264588
1653	94255993 (3305, 3306)	Novel Protein sim. GBank gj13776054[emb]CAA06273] - (AJ004889) Tapasin [Gallus gallus]	Contains protein domain (PF00047) - immunoglobulin domain	glycoprotein	
1654	79758471 (3307, 3308)			UNCLASSIFIED	

1655	66689346 (3308, 3310)	Novel Protein sim. GBank gij3355717[emb]CAA73498] - (Y13053) senyI-RNA synthetase [Zea mays]		synthase	52844507, 35686286, 22276996, 22276999, 28331624, 28331625, 29331628, 33659870, 264908, 52644045, 264511, 264910, 52646317, 264288, 52644229, 33657023, 33657109, 52644332, 264557, 56162323, 56526488, 80432113
1656	79862297 (3311, 3312)	Novel Protein sim. GBank gij1890141[dbj]BAA18947] - (D63206) P24 protein [Mus musculus]		UNCLASSIFIED	28331622, 28331624, 28331625, 264563
1657	67771994 (3313, 3314)	Novel Protein sim. GBank gij4557645[ref]NP_001524.1[phn]RP - heterogeneous nuclear ribonucleoprotein L	Contains protein domain (PF00076) - RNA recognition motif. (a.k.a. RRM, RBD, or RNP domain)	dna_ma_bind	29331627, 265009, 21906766, 21906767, 285020, 265022, 33657108, 264638, 56526488, 264462
1658	67773776 (3315, 3318)	Novel Protein sim. GBank gij3877072[emb]CAA67060] - (Z46837) similarity with ribosomal protein L21 [Caenorhabditis elegans]	Contains protein domain (PF00829) - Ribosomal prokaryotic L21 protein	UNCLASSIFIED	52848365, 35686286, 22276996, 22276997, 22276998, 22276999, 264259, 29331822, 29331624, 29331825, 29331826, 29331827, 28331628, 29148498, 264905, 284908, 52644045, 265006, 60433356, 264757, 60433438, 21906754, 265011, 18108351, 284448, 264369, 264288, 264766, 264768, 21906765, 21906767, 21906768, 21906769, 29148629, 265021, 265022, 18106362, 263969, 263971, 18108374, 35686423, 18108383, 22279000, 264482
1659	68230101 (3317, 3318)	Novel Protein sim. GBank gij539218[pi]S36036 - hypothetical protein YKL201c - yeast (Saccharomyces cerevisiae)		UNCLASSIFIED	52846317, 21906766, 21906767, 21906768, 67168516, 22276996, 265020, 22276999, 67168558, 264603, 285017, 264631, 285016, 265019, 22279002, 264462, 264635, 264565
1660	94315313 (3319, 3320)	Novel Protein sim. GBank gij2497012[sp]Q10010]YSV4_CAEEL - HYPOTHETICAL 26.8 KD PROTEIN T19C3.4 IN CHROMOSOME III		UNCLASSIFIED	264486, 35686286, 264259, 35686032, 264506, 264509, 264905, 264906, 264907, 264908, 264909, 264510, 264511, 265006, 265007, 264512, 265008, 264910, 264592, 264598, 265010, 264600, 264802, 265017, 265018, 264805, 264760, 264764, 264288, 264766, 264666, 264766, 264789, 264689, 21906766, 35685917, 264690, 33657023, 264693, 33657109, 264629, 35686423, 35685855, 264634, 264635, 264555, 264636, 264637, 264556, 264638, 264638, 264559, 18108385, 18108388, 284563, 284483, 284564, 284565, 264566, 264486, 264567

1661	94234071 (3321, 3322)	Novel Protein sim. GBank gi 4759100 ref NP_004759.1 pSFRS - splicing factor, arginine/serine-rich 11	Contains protein domain (PF00076) - RNA recognition motif (a.k.a. RRM, RBD, or RNP domain)	UNCLASSIFIED	264466, 22276998, 264259, 29331624, 29331828, 29331627, 29331628, 264509, 66712502, 29331830, 264906, 52844045, 265007, 264512, 60433356, 60433438, 55812038, 21906754, 265019, 284446, 264766, 264766, 264769, 21906766, 21906769, 265020, 33657023, 33657109, 65274791, 67166516, 264482, 264563, 264564, 264565, 264567
1662	94135172 (3323, 3324)	Novel Protein sim. GBank gi 1730502 sp P52675 PF27_MOUSE - TRANSMEMBRANE PROTEIN PF27			16108392, 29331622, 29331828, 20261100, 264106, 265006, 265007, 265008, 18108346, 21906766, 18108365, 18108368, 16108374, 83373044, 18108385
1663	94217146 (3325, 3326)	Novel Protein sim. GBank gi 4684136 emb CAB43275.1 - (AL050107) hypothetical protein [Homo sapiens]	Contains protein domain (PF00397) - WW domain	kinase	52645156, 56162575, 22276994, 22276995, 35696288, 22276996, 58994075, 22276997, 22278998, 22278999, 264258, 29331822, 29331626, 29331627, 29331628, 33656970, 29331830, 264906, 56162435, 264511, 60433356, 33657402, 33108954, 87166474, 87166559, 265017, 265018, 264605, 16108351, 264764, 264268, 264766, 264766, 21906765, 21906768, 21906767, 21906768, 21906769, 265021, 265022, 264691, 33657023, 264693, 263967, 33657109, 264630, 52644332, 63373044, 87168516, 60432113, 22279000
1664	94234076 (3327, 3328)	Novel Protein sim. GBank gi 3043892 dbj BAA25510 - (AB011156) KIA0584 protein [Homo sapiens]		UNCLASSIFIED	264486, 263994, 35696266, 29331824, 35696052, 264508, 264509, 264905, 264906, 264907, 264908, 264909, 264510, 264511, 265009, 264910, 60170631, 264591, 264592, 264595, 87188474, 265011, 264600, 264601, 264604, 264605, 264780, 264762, 18108351, 264681, 264682, 264763, 264663, 264764, 264266, 264684, 264766, 264687, 264766, 264769, 21906764, 21906765, 21906767, 264690, 264691, 264692, 33657109, 35695917, 265021, 264534, 60170615, 33657162, 264628, 18108370, 264629, 35696423, 35695855, 264634, 264635, 264555, 264636, 264637, 264638, 264639, 264556, 83373044, 87168518, 264563, 264586, 264486

1665	91226952 (3329, 3330)	Novel Protein sim. GBank gll1083506[pr][S50065 - slalodhesin - mouse	Contains protein domain (PF00047) - Immunoglobulin domain	immunoglob	264488, 29331628, 29331628, 284509, 284908, 284907, 284908, 284510, 284511, 284910, 284592, 284593, 284595, 284758, 284598, 284600, 284780, 284782, 284784, 284786, 284788, 284629, 284630, 284634, 284636, 83373044, 284564, 284568, 284587, 284488
1666	95358160 (3331, 3332)	Novel Protein sim. GBank gll3913431[sp][Ox2843][DX8_SCHPO - PUTATIVE PRE-MRNA SPLICING FACTOR ATP-DEPENDENT RNA HELICASE SPAC10F6.02C	Contains protein domain (PF00575) - S1 RNA binding domain	helicase	56984075, 22278999, 284259, 29331824, 29331628, 29331827, 29146488, 285009, 33109954, 87188559, 285019, 284288, 284688, 21908787, 21908789, 284681, 33857182, 18108370, 18108374, 18108385, 22279002
1667	91228855 (3333, 3334)	Novel Protein sim. GBank gll5689535[db][BAA83051.1] - (AB029022) KIAA1099 protein [Homo sapiens]	Contains protein domain (PF01412) - Putative GTP-ase activating protein for Arf	transport	264259, 29331622, 29331626, 284905, 284908, 284908, 284510, 285009, 284595, 284758, 265011, 67168559, 265017, 265018, 265019, 284448, 284788, 284886, 21908785, 21908787, 21908789, 285020, 265021, 60170615, 284690, 284692, 284693, 18108368, 18108370, 283972, 55810784, 284555, 83373044, 60432113, 22279000, 22279002
1668	88095135 (3335, 3336)	Novel Protein sim. GBank gll2076894[gb][AAB53983.1] - (AF002197) short region of weak similarity to protein kinase C; contains similarity to Pfam domain PF00130 (DAG_PE-bind). Score=10.0, E-value=0.0034. N=1 [Caenorhabditis elegans]		kinase	66714117, 284508, 284509, 284906, 284907, 284908, 284511, 284910, 284784, 284887, 284688, 33857109, 35696423, 35695855, 264632
1669	91227846 (3337, 3338)	Novel Protein sim. GBank gll3875371[emb][CAA85414.1] - (Z36848) contains a valine and arginine rich domain, possesses weak similarity with the RNA binding domains from RNA splicing factor U2AF 65 KD subunit; cDNA EST EMBL:D64658 comes from this gene; cDNA EST EMBL:D68829 comes fr...		UNCLASSIFIED	29331825, 33109954, 284369, 284787, 264689, 33857109, 63373044
1670	67826009 (3339, 3340)			UNCLASSIFIED	264259, 29331624, 29331627, 80433438, 265022, 284638
1671	67346372 (3341, 3342)	Novel Protein sim. GBank gll482451[sp][P3424][KKK1_YEAST - PROBABLE SERINE/THREONINE-PROTEIN KINASE YKL101W	Contains protein domain (PF00069) - Eukaryotic protein kinase domain	kinase	35998286, 22278997, 29331625, 284909, 21908754, 265017, 265018, 265019, 284882, 284683, 284786, 284888, 21908786, 21908787, 21908789, 21908789, 284681, 284555, 284558, 22279000, 284588
1672	86291834 (3343, 3344)	Novel Protein sim. GBank gll1814270 (U74586) - double-stranded RNA specific adenosine deaminase [Rattus norvegicus]		deaminase	264906, 284908, 284632, 18108381

1673	68095137 (3345, 3346)	Novel Protein sim. GBank gi 2076894 gb AAB53983.1 - (AF002197) short region of weak similarity to protein kinase C; contains similarity to Pfam domain PF00130 (DAG_PE-bind). Score=10.0, E-value=0.0034, N=1 [Caenorhabditis elegans]	Contains protein domain (PF00130) - Phorbol esters/diacylglycerol binding domain (C1 domain)	kinase	264488, 264569, 16108394, 56994075, 22278996, 264259, 35696052, 264508, 264905, 264509, 264906, 264907, 264908, 264909, 264510, 264910, 60170831, 264592, 264594, 264595, 264758, 264801, 264760, 264762, 264683, 264764, 264288, 264768, 264886, 264758, 264687, 264769, 264889, 264890, 33657023, 264692, 264693, 33657109, 264628, 264829, 16108374, 35696423, 35695855, 264631, 264632, 264634, 264635, 264637, 264558, 264638, 264639, 264563, 264482, 264564, 264565, 264568, 264587, 264488
1674	86258028 (3347, 3348)	Novel Protein sim. GBank gi 5262467 emb CAB45693.1 - (AL080082) hypothetical protein [Homo sapiens]		kinase	29331822, 29331824, 264908, 52644045, 60433358, 67168559, 264446, 264288, 264686, 264691
1675	67606466 (3349, 3350)	Novel Protein sim. GBank gi 3128366 (AF010486) - 50S ribosomal protein 19 [Rhodobacter capsulatus]		UNCLASSIFIED	56161888, 35696286, 22278997, 22278998, 264259, 29331824, 29331827, 35696052, 66712502, 264764, 264288, 264688, 264687, 35695917, 265020, 264890, 264893, 35695763, 16108370, 35696423, 35695855, 264637, 264639, 16108385, 264564
1676	95358086 (3351, 3352)	Novel Protein sim. GBank gi 164065 gb AAD05327 - (AF111091) latrophilin 3 splice variant btaf [Bos taurus]		UNCLASSIFIED	264259, 29331827, 29331828, 264106, 264907, 265009, 264600, 265019, 264288, 21908765, 265020, 265022, 35695855, 63373044, 16108385
1677	67406567 (3353, 3354)	Novel Protein sim. GBank gi 3327046 db BAA31591 - (AB014516) KIAA0616 protein [Homo sapiens]		UNCLASSIFIED	264908
1678	86866829 (3355, 3356)	Novel Protein sim. GBank gi 550432 (U08469) - 3-methylcrotonyl-CoA carboxylase, biotin-carrier domain [Glycine max]	Contains protein domain (PF00289) - Carboxymethyl-phosphate synthase (CPase)	UNCLASSIFIED	29331824, 264102
1679	91214108 (3357, 3358)	Novel Protein sim. GBank gi 550432 (U08469) - 3-methylcrotonyl-CoA carboxylase, biotin-carrier domain [Glycine max]		carboxylase	264468, 16108392, 16108396, 35696288, 29331824, 265008, 285007, 285008, 285009, 16108346, 265011, 16108351, 264683, 16108354, 16108358, 16108359, 21908765, 29148827, 29148829, 264690, 16108381, 16108382, 16108384, 16108385, 16108386, 264635, 16108379, 35696423, 35695855, 264635, 16108381, 16108382, 16108383, 16108384, 16108385, 16108386
1680	81005372 (3359, 3360)	Novel Protein sim. GBank gi 2394478 (AF024500) - No definition line found [Caenorhabditis elegans]		transport	55274572, 22278994, 22278999, 66714117, 29331827, 56162435, 21908754, 265016, 264286, 21906769
1681	94324150 (3361, 3362)	Novel Protein sim. GBank gi 5689537 db BAA63052.1 - (AB029023) KIAA1100 protein [Homo sapiens]		UNCLASSIFIED	22278996, 29331822, 264908, 264593, 264604, 265019, 264683, 55811957, 264690, 33657023, 35696423, 63373044, 264563
1682	86042710 (3363, 3364)			UNCLASSIFIED	264909, 285017, 284605

1683	84316213 (3365, 3366)	Novel Protein sim. GBank g 5031717 refNP_005704.1 pGPBP - goodpasture antigen-binding protein	Contains protein domain (PF01852) - UNCLASSIFIED START domain	263994, 35696286, 35696052, 264508, 264509, 264905, 264908, 264907, 264908, 264909, 264511, 265008, 265007, 265009, 264910, 264593, 264758, 265010, 265011, 264760, 264761, 264762, 264763, 264764, 264369, 264766, 264768, 35695917, 264692, 33657109, 35696423, 35695855, 264834, 264635, 264636, 264638, 264639, 83373044, 264486
1684	80063409 (3367, 3368)			264583, 264586
1685	94323182 (3389, 3370)	Novel Protein sim. GBank g 1255371 (U53147) - coded for by C. elegans cDNA yk34a9.5; coded for by C. elegans cDNA yk34a9.3; Similar to guanylate kinase. [Caenorhabditis elegans]	Contains protein domain (PF00625) - UNCLASSIFIED kinase Guanylate kinase	60424179, 52646842, 22278994, 35696286, 22278998, 264259, 52645080, 29331824, 29331826, 265007, 33857084, 265018, 264681, 264448, 264683, 264369, 264689, 21908765, 21908787, 21908788, 21908789, 265021, 264682, 65274820, 33657109, 27486262, 264635, 52644332, 58182323, 22279000
1686	87620710 (3371, 3372)	Novel Protein sim. GBank g 2244707 db BAA31115.1 - (AB005287) thrombospondin 1 [Bos taurus]	UNCLASSIFIED	56182575, 264259, 264508, 264905, 264509, 264907, 264908, 264510, 264511, 265006, 264512, 265008, 264910, 264756, 265010, 265011, 264605, 18108351, 264764, 264766, 18108357, 264768, 18108362, 264628, 264630, 264631, 264834, 264835, 264637, 264638, 264639, 264585, 264466, 264587
1687	94719400 (3373, 3374)	Novel Protein sim. GBank g 4680679 gb AAD27729.1 AF13295 - (AF132954) CGI-20 protein [Homo sapiens]	UNCLASSIFIED	35695917, 264906, 264907, 264906, 264510, 265008, 265007, 264910, 264558, 18108381, 18108383, 285011
1688	82158442 (3375, 3376)		UNCLASSIFIED	264905, 264910, 264760, 264629, 264555
1689	94325049 (3377, 3378)	Novel Protein sim. GBank g 4240193 db BAA74875.1 - (AB020659) KIAA0852 protein [Homo sapiens]	UNCLASSIFIED	264569, 35696286, 22278999, 264508, 264908, 264909, 56182435, 33657402, 80433438, 55812038, 265017, 265018, 264446, 264764, 264288, 264686, 29148629, 35695917, 265020, 285021, 263972, 18108374, 65274791, 83373044, 264089
1690	83255346 (3379, 3380)	Novel Protein sim. GBank g 3800736 (AF031572) - seven-pass transmembrane receptor precursor [Mus musculus]	Contains protein domain (PF00026) - cadherin	264369, 21906766, 264692, 264639, 87188518
1691	88095223 (3381, 3382)	Novel Protein sim. GBank g 2773208 (AF039713) - No definition line found [Caenorhabditis elegans]		264786, 33657109, 29331827, 29148629, 264510, 264106, 264910, 264109, 264508, 60170831, 264583, 264905, 264584, 264691, 264837, 264828, 264907, 264908, 33657023, 264587, 264768, 263974
1692	86106709 (3383, 3384)			264106
1693	87012775 (3385, 3386)	Novel Protein sim. GBank g 121271 sp P02207 GLB_LAMFL - GLOBIN	Contains protein domain (PF00042) - UNCLASSIFIED Globin	29331826, 264508, 264905, 264907, 264595, 265010, 265011, 21908788, 33657023, 264629, 263978, 264558

1694	94208188 (3387, 3388)	Novel Protein sim. GBank gi5453932[refNP_006225.1]pPOLR - polymerase (RNA) II (DNA directed) polypeptide J (13.3kD)	Contains protein domain (PF01193) - RNA polymerases L / 13 to 18 kDa subunit	mapolymerase	35698286, 22278998, 22278998, 22278998, 284259, 29331822, 29331825, 29331828, 29331828, 35696052, 29146499, 284905, 284908, 52644045, 264511, 285008, 265007, 265009, 264592, 60433358, 21908754, 285010, 285011, 18108351, 264763, 264682, 264448, 264683, 284288, 284788, 264889, 21908788, 80170815, 284691, 284692, 284693, 18108370, 18108374, 283978, 35896423, 35695855, 264558, 18108381, 18108385, 87188518, 284482, 284488 284634
1695	94718325 (3389, 3390)	Novel Protein sim. GBank gi4860879[gbAAD27729.1]AF13295 - (AF132954) CGI-20 protein [Homo sapiens]	UNCLASSIFIED	UNCLASSIFIED	
1698	87824038 (3391, 3392)	Novel Protein sim. GBank gi4220517[embCAA22890] - (AL035356) hypothetical protein [Arabidopsis thaliana]			22278897, 264259, 285010, 18108351, 264764, 21908766, 18108370
1697	85740963 (3393, 3394)	Novel Protein sim. GBank gi505662 (U10362) - GP36b glycoprotein [Homo sapiens]		glycoprotein	284682
1698	87445285 (3395, 3396)	Novel Protein sim. GBank gi5032031[gbAAD38411.1]AF15573 - (AF155739) axotrophin [Mus musculus]			56994075, 22278998, 22278999, 284509, 33657402, 264758, 87188474, 87188559, 285017, 265018, 264448, 264687, 29148827, 21908769, 29148629, 285020, 285022, 33657023, 264558, 87188518, 22279002 35898288, 284635
1699	87424783 (3397, 3398)	Novel Protein sim. GBank gi543344[prj]S41647 - zinc finger 5 protein - mouse	Contains protein domain (PF00058) - Zinc finger, C2H2 type	nuc_rept	28331824, 52844045, 265008, 265009, 263969, 263971
1700	87859181 (3399, 3400)	Novel Protein sim. GBank gi3877439[emb]CAA98652] - (Z72510) similarity to yeast UTR3 protein (Swiss Prot accession number P21374); cDNA EST EMBL:D72622 comes from this gene; cDNA EST EMBL:D75763 comes from this gene; cDNA EST YK274e3.3 comes from this gene; cDNA EST YK274e3.5 c...		MHC	284092, 284110, 263977 22278995, 22278997, 284092, 29146498, 29146499, 284107, 264508, 264907, 284110, 264112, 285009, 60170831, 21908754, 285011, 285017, 264782, 18108351, 284288, 21908765, 35695917, 265021, 60170815, 263987, 33657109, 18108370, 263972, 263974, 18108374, 283978, 35695855, 264555, 263981, 80170394, 18108385, 56528486, 87188518, 60432113
1703	79588651 (3405, 3406)	Novel Protein sim. GBank gi451544 (U04287) - proline-rich cell wall protein [Gossypium barbadense]		UNCLASSIFIED	284909, 265017, 264628, 264629, 284838
1704	86622979 (3407, 3408)	Novel Protein sim. GBank gi1263289 (U47856) - fibrin-4 [Araucaria diadematus]		UNCLASSIFIED	264389
1705	87785175 (3409, 3410)	Novel Protein sim. GBank gi4519621[gb]BA75870.1] - (AB017614) OASIS protein [Mus musculus]			284569, 35696288, 284807, 285010, 264687, 284788, 264892, 264893, 264836, 264588 22278998, 22278998, 264259, 264509, 285018, 284764, 264885, 284886, 21908788, 21908769, 265022, 264891, 284558, 22279000
1706	87790987 (3411, 3412)	Novel Protein sim. GBank gi3123034[ppQ15011]Y025_HUMAN - HYPOTHETICAL PROTEIN KIAA0025			

1707	86041230 (3413, 3414)	Novel Protein sim. GBank gi 4321664 gb AAD15791 - (AF055470) ZNF258 [Homo sapiens]		UNCLASSIFIED	16106396, 22278997, 264259, 29147620, 29331626, 29148498, 264905, 264906, 285008, 264593, 264595, 264758, 264598, 265016, 264760, 16108351, 264764, 264766, 264689, 264693, 16106370, 35696423, 55811576, 264556, 67168518, 60432113, 264567
1708	91220519 (3415, 3416)	Novel Protein sim. GBank gi 5174591 ref NP_005947.1 pMTFH - 5,10-methylenetetrahydrofolate dehydrogenase, 5,10-methylenetetrahydrofolate cyclohydrolase, 10-formyltetrahydrofolate synthetase	Contains protein domain (PF01268) - Formate-tetrahydrofolate ligase	synthase	56162575, 22278996, 56994075, 264259, 29331622, 29331624, 29331626, 29331627, 29331628, 29146498, 29146499, 29331630, 265009, 60170831, 33657402, 33109854, 67168559, 265019, 16108351, 264446, 21908765, 21906767, 21906768, 29146627, 29146629, 29148784, 60170615, 52644150, 33657023, 33657109, 16106374, 55611576, 264559, 16108385, 22279000, 264563
1709	60222583 (3417, 3418)			UNCLASSIFIED	264107, 55611957, 263974, 263976, 263977, 263981
1710	20754572 (3419, 3420)			UNCLASSIFIED	264556
1711	91013729 (3421, 3422)	Novel Protein sim. GBank gi 5031735 ref NP_005760.1 pHEC - N-acetylglucosamine 6-O-sulfotransferase		sulfotransferase	65274572, 29331824, 29331826, 264768, 60431526, 35696423, 60432113, 264563
1712	95330164 (3423, 3424)	Novel Protein sim. GBank gi 5454168 ref NP_008453.1 pXAP4 - HBV associated factor	Contains protein domain (PF00641) - Zn-finger in Ran binding protein and others.	kinase	56984075, 264093, 264259, 29331822, 264099, 29331824, 29331827, 264107, 264110, 264511, 264592, 265011, 265018, 264683, 264686, 264689, 265020, 33657023, 263987, 33657109, 263974, 35696423, 35695855, 264630, 264636, 264558, 264566
1713	94143453 (3425, 3426)	Novel Protein sim. GBank gi 160409 (M89183) - mature-parasite-infected erythrocyte surface antigen [Plasmodium falciparum]	Contains protein domain (PF00643) - B-box zinc finger.	UNCLASSIFIED	22276895, 264508, 264758, 16108351, 16108370, 263974, 16106374, 264634, 56162323, 83373044, 60432113
1714	87420046 (3427, 3428)				22276897, 264757, 21906765, 265020, 265021, 264692, 56528466
1715	84260257 (3429, 3430)	Novel Protein sim. GBank gi 5689537 dbj BA463052.1 - (AB029023) KIAA1100 protein [Homo sapiens]		UNCLASSIFIED	264508, 264905, 264906, 264907, 264908, 264909, 264910, 264591, 265011, 264766, 264768, 264769, 264691, 264692, 264632, 264634, 264635, 264636, 264637, 264556, 264639, 264584
1716	87400449 (3431, 3432)	Novel Protein sim. GBank gi 4589468 dbj BA476761.1 - (AB012809) mBOCT [Mus musculus]		transport	56162575, 29331624, 60432269, 264109, 264909, 265007, 264800, 265019, 264686, 265020, 264893, 55611576, 264556, 60432113, 22278902

1717	67563223 (3433, 3434)	Novel Protein sim. GBank glij2755411emh[CAA74749] - (Y14391) GTP-binding protein [Homo sapiens]		UNCLASSIFIED	264569, 264258, 29331825, 29331828, 29331829, 35696052, 264508, 264905, 264907, 264908, 264909, 264512, 285009, 264910, 264592, 264595, 264758, 264759, 265017, 264681, 264784, 264766, 264688, 18108357, 35695917, 264690, 264692, 264683, 264628, 264629, 35696423, 264630, 264631, 264635, 264636, 16108380, 264638, 264639, 18108388, 18108391
1718	87032828 (3435, 3436)	Novel Protein sim. GBank glij26332825pqi14999Y076_HUMAN - HYPOTHETICAL PROTEIN KIAA0076 (HA0936)		UNCLASSIFIED	265011, 264881, 264882, 264684, 264688, 264689, 21908765, 265021, 264691, 33657023, 264693, 18108370, 35695855, 264832, 264634, 264636, 16106386, 22279002
1719	94315259 (3437, 3438)	Novel Protein sim. GBank glij4505197ie[MP_003473.1]pML2 - myeloidlymphoid or mixed-lineage leukemia 2		UNCLASSIFIED	16108396, 65274572, 35696286, 22278997, 60432049, 58182181, 68714117, 60432269, 29331826, 35696052, 29331828, 264906, 29331830, 58182435, 264582, 60431735, 60433438, 5812036, 264759, 265010, 264600, 264601, 265017, 264448, 264764, 264288, 264769, 21908766, 21906769, 55611857, 265020, 265021, 52844150, 33657023, 33657109, 33657182, 27486262, 33657348, 35695763, 18108370, 60431528, 18108374, 35696423, 55811576, 35695855, 264631, 56162323, 264538, 264584, 264486, 56182575, 22278998, 264259, 29331824, 60432289, 29331827, 35696052, 264508, 264905, 264906, 264807, 264808, 264809, 264511, 264910, 264758, 21906754, 265011, 264801, 264760, 264762, 264288, 264766, 264866, 18108357, 264689, 21906765, 55611957, 264693, 20281149, 264629, 18108374, 55811576, 65274791, 264630, 20281071, 264834, 264635, 264838, 264637, 264536, 264638, 264639, 56162323, 87169518
1720	84653063 (3439, 3440)	Novel Protein sim. GBank glij2129478p[S51939 - chitinase (EC 3.2.1.14) precursor - beet		UNCLASSIFIED	22278994, 22278999, 29331822, 285006, 265007, 265008, 55812036, 21906754, 60174639, 265011, 87168559, 18108351, 16108354, 21908765, 21908766, 21908768, 21906769, 285020, 33657109, 18108370, 16108374, 264556, 60170394, 63373044, 16108385, 264488
1721	91722268 (3441, 3442)	Novel Protein sim. GBank glij4886461emh[CAB3381.1] - (AL050280) hypothetical protein [Homo sapiens]		UNCLASSIFIED	56994075, 28331824, 29331828, 265008, 18108351, 21908768, 285020, 33657023, 18108374, 83373044
1722	94134549 (3443, 3444)	Novel Protein sim. GBank glij5688375idh[BAA82986.1] - (AB030644) tudor repeat associator with PCTAIRE 2 [Rattus norvegicus]	Contains protein domain (PF00567) - Kinase Tudor domain		

1723	95358181 (3445, 3446)	Novel Protein sim. GBank g1426862/gb1A020631 - (AF126062) Art-like 2 binding protein BART1 [Homo sapiens]		UNCLASSIFIED	264488, 264687, 264769, 21906767, 21906768, 58182575, 55811957, 22278997, 22278998, 285020, 264259, 264892, 33857023, 29331822, 29331824, 29331825, 80432289, 33657182, 33656970, 33657348, 29148499, 264508, 264907, 18108370, 264829, 264808, 264909, 18108374, 55811578, 264510, 285008, 264511, 265007, 264910, 264832, 264591, 80432228, 264592, 60433356, 264594, 60433438, 264595, 83373044, 55812038, 33109954, 33657084, 87188518, 87188474, 285010, 285011, 87188559, 264600, 60432113, 264604, 265019, 264563, 284448, 264682, 264588, 264784, 264286, 284587, 264488, 264369, 264768
1724	87713806 (3447, 3448)	Novel Protein sim. GBank g12340162 (AF005083) - dsRBP-ZF a [Xenopus laevis]	Contains protein domain (PF00098) - Zinc finger, C2H2 type	UNCLASSIFIED	264905, 18108359, 284693, 264828, 264631, 264638, 264555, 264556, 264558, 264559, 35696286, 264259, 29331822, 35696052, 284508, 264509, 264905, 264908, 284907, 264808, 264809, 264910, 265009, 264891, 264601, 264780, 18108351, 264881, 264764, 264288, 264788, 264788, 21906769, 35685917, 264628, 35688423, 264630, 264831, 264632, 264635, 264638, 264838, 87188518, 264588
1725	85855191 (3449, 3450)	Novel Protein sim. GBank g13152662 (AF064604) - KEO3 protein [Homo sapiens]	Contains protein domain (PF00023) - Ank repeat	homeobox	29146498, 264683, 264689
1726	85754235 (3451, 3452)	Novel Protein sim. GBank g14689348/gb1A027881.1/AF13256 - (AF132582)		UNCLASSIFIED	264905, 285011, 264889, 21908766
1727	85286382 (3453, 3454)	Novel Protein sim. GBank g14689348/gb1A027881.1/AF13256 - (AF132582)		UNCLASSIFIED	

1728	95340515 (3455, 3456)	Novel Protein sim. GBank gll4406548 gblAAD200271 - (AF131738) Unknown [Homo sapiens]		UNCLASSIFIED	60424170, 16106397, 56162575, 22276995, 56994075, 35698286, 22276997, 22276998, 22277999, 264094, 60432049, 264259, 29331822, 29331824, 56182181, 29331825, 60432289, 29331828, 29331827, 35696052, 264905, 264906, 264907, 29331830, 66712502, 264906, 56162435, 264511, 265008, 265009, 60432229, 60433356, 33657402, 60433436, 264759, 21906754, 67166474, 265010, 265011, 87166559, 265017, 265016, 265019, 55611150, 264681, 264446, 264682, 264783, 264663, 264286, 264684, 264369, 264685, 264766, 264687, 264769, 21906764, 264669, 21906765, 21906766, 21906767, 21906768, 35695917, 265020, 265021, 265022, 264535, 264691, 264692, 33657023, 264693, 33657109, 18108370, 264628, 263972, 264629, 18108374, 18106376, 55810764, 65274791, 35695855, 264631, 264634, 264635, 60431850, 264836, 264636, 60170394, 264638, 63373044, 56526466, 67166518, 60432113, 22279000, 22279002, 264564, 264566
1729	91227948 (3457, 3458)	Novel Protein sim. GBank gll854065 emb CAA563371 - (X63413) U66 [Human hemepusvus 6]		UNCLASSIFIED	264906, 264907, 264906, 264511, 264555, 63373044, 264596, 264566
1730	85463474 (3459, 3460)			UNCLASSIFIED	29331822, 29331825, 29331826, 264907, 264908, 264909, 265011, 264764, 264629
1731	86266066 (3461, 3462)	Novel Protein sim. GBank gll631600 p S47084 - hypothetical protein - rabbl		UNCLASSIFIED	52646842, 264807, 264908, 56182435, 55611366, 67166559, 265016, 265019, 264760, 52644229, 55811576
1732	91216878 (3463, 3464)	Novel Protein sim. GBank gll4240231 db BAV4684.1 - (AB020676) KIAA0871 protein [Homo sapiens]		struct	56162575, 29331822, 29331824, 29331827, 66712502, 264591, 33657402, 60433356, 265019, 21906768, 21906769, 35695917, 265020, 265021, 264638, 56162323
1733	67617178 (3465, 3466)	Novel Protein sim. GBank gll157576 (U70674) - m-Humb [Mus musculus]	Contains protein domain (PT00640) - synthase Phosphotyrosine interaction domain (PTB/PID).		264807, 264810, 33657402, 265010, 264661, 264683, 264684, 264686, 264769, 264691, 264692, 264693, 264628, 264636, 264556
1734	67785261 (3467, 3468)				264693

1735	88318638 (3469, 3470)	Novel Protein sim. GBank gi14836807 gb AA030588.1 AF14679 - (AF146793) PF127 (Mus musculus)	Contains protein domain (PF01169) - Uncharacterized protein family UPF0016		264488, 18108394, 18108398, 22278998, 60432049, 264259, 29331822, 264908, 265006, 265007, 265008, 265009, 264591, 33657402, 265010, 265011, 87168559, 265017, 18108351, 264882, 18108354, 264769, 264889, 21906765, 21906768, 21906769, 29148629, 29148784, 265021, 265022, 52644150, 18108384, 18108385, 33657109, 18108370, 18108374, 18108380, 18108385, 87168518, 264583, 18108390, 264490, 264259, 66714117, 88712502, 58182435, 285006, 285008, 264910, 60433358, 87168559, 265017, 265018, 18108351, 285020, 265022, 33657023, 33657109, 264555, 60431850, 264637, 60170394, 264558, 264639, 264584, 29331822, 29331830, 264591, 265011, 285018, 285019, 22278902, 264908, 264909, 285008, 264910, 264566, 33696052, 264603, 264557
1736	85362884 (3471, 3472)	Novel Protein sim. GBank gi4885847 en NP_005472.1 PTRAP - thyroid hormone receptor-associated protein complex component		UNCLASSIFIED	
1737	88165549 (3473, 3474)	Novel Protein sim. GBank gi12143607 p S68695 - BK protein - rat	Contains protein domain (PF00169) - C2 domain	kinase	
1738	85788811 (3475, 3478)	Novel Protein sim. GBank gi1222594 t emb CAA69714 - (Y08460) Mdes protein [Mus musculus]	Contains protein domain (PF00441) - Acyl-CoA dehydrogenase	UNCLASSIFIED	
1739	87326576 (3477, 3478)			UNCLASSIFIED	
1740	83592839 (3479, 3480)	Novel Protein sim. GBank gi14809 emb CAA44309 - (X62452) YCR601 [Saccharomyces cerevisiae]		traffic	264604, 21906764, 18108384, 264629, 35695855, 264636
1741	95010100 (3481, 3482)	Novel Protein sim. GBank gi1483898 p AAD31695.1 AF13042 - (AF130420) serine protease-like protein isoform [Homo sapiens]	Contains protein domain (PF01383) - FYVE zinc finger	UNCLASSIFIED	80432229, 29331827, 264509, 265009, 60432229, 264759, 265017, 264767, 264688, 264689, 21906768, 265020, 33657109, 264908, 264910, 264758, 265011, 264631, 264638, 264588
1742	85788814 (3483, 3484)	Novel Protein sim. GBank gi14505183 en NP_003687.1 p MLD - membrane fatty acid (lipid) desaturase		UNCLASSIFIED	
1743	8896475 (3485, 3486)	Novel Protein sim. GBank gi172832 sp P39189 ALU2_HUMAN - III ALU SUBFAMILY SB WARNING ENTRY III		UNCLASSIFIED	265017, 265020, 284692, 22278988, 264508, 264907
1744	91224003 (3487, 3488)			UNCLASSIFIED	
1745	20280075 (3489, 3490)	Novel Protein sim. GBank gi1731758 sp P38873 HYH8_YEAST - HYPOTHEICAL 175.8 KD PROTEIN IN GND1-K1 INTERGENIC REGION (Drosophila melanogaster)	Contains protein domain (PF00400) - WD domain, G-beta repeat	UNCLASSIFIED	264558 65274791, 264638, 264559
1746	94328110 (3491, 3492)			transcription factor	52646842, 29331824, 29331825, 52644045, 56182435, 265007, 52648317, 87168474, 265018, 18108351, 264380, 264769, 264689, 33695917, 60170815, 264681, 33657023, 65274791, 35695855, 60170394, 22279000, 264482
1747	94324333 (3493, 3494)				

1748	88003580 (3495, 3498)	Novel Protein sim. GBank gi14504511 pe NP_001530.1 p HSJ2 - heat shock protein, DNA-like 2	Contains protein domain (PF00864) - eph DnaJ central domain (4 repeats)		264489, 5618257, 29331824, 56182435, 284112, 2185007, 265019, 264784, 21908788, 265020, 264681, 55811578, 284835, 284555, 264558, 284557, 264559
1749	83363081 (3487, 3498)	Novel Protein sim. GBank gi15650780 p AAC45948.1 AF15196 - (AF15196) RGS protein RCS-17 (Gallus gallus)	Contains protein domain (PF00815) - oncogene Regulator of G protein signaling domain		264106
1750	94321664 (3498, 3500)	Novel Protein sim. GBank gi14996894 p AAC28444.2 - (AF085184) hyperpolarization-activated, cyclic nucleotide- gated channel 2 (Homo sapiens)			33657402, 264288, 52644150, 283974, 83373044
1751	83373058 (3501, 3502)	Novel Protein sim. GBank gi12760181 p BAA24184 - (AB010054) outer arm dynein light chain 2 (Anthodaris crassipal)	Contains protein domain (PF00560) - ATPase, associated Leucine Rich Repeat		265010, 264369
1752	86466530 (3503, 3504)	Novel Protein sim. GBank gi13915482 p P74346 YGC29, SYNY3 - HYPOTHETICAL 36.0 KD PROTEIN SLR1828	Contains protein domain (PF00849) - deaminase RNA pseudouridylation synthase		264510, 264593, 264682, 21906765, 18108370
1753	84235159 (3505, 3508)	Novel Protein sim. GBank gi12852836 (AF007155) - unknown (Homo sapiens)	Contains protein domain (PF01553) - phosphatase Acyltransferase		56934075, 22278996, 264908, 60170831, 284882, 264784, 264389, 284288, 284885, 284687, 21906766, 264692, 264693, 85274820, 85274791, 35895855, 284637, 284584
1754	88095323 (3507, 3508)	Novel Protein sim. GBank gi1731421 p P39881 YEH_YEAST - HYPOTHETICAL 53.3 KD PROTEIN IN HXT8-CAN1 INTERGENIC REGION	transport		264498, 35896286, 264509, 264906, 264907, 264908, 264909, 264511, 264910, 264591, 33657402, 264594, 264757, 264758, 264600, 284804, 264782, 264783, 264685, 264768, 264891, 264628, 35896423, 264832, 264634, 284837, 264838, 263981, 264558, 284639, 284563, 264564, 264585, 264568, 264587, 264686
1755	79470282 (3509, 3510)	Novel Protein sim. GBank gi11176422 (U43194) - tropilin (Mus musculus)	UNCLASSIFIED		5264642, 22278994, 22278995, 56994075, 22278998, 22278997, 264259, 29331822, 60432288, 29331827, 33658970, 265006, 265009, 60432229, 60433356, 80433438, 33109954, 21808754, 265017, 265018, 285019, 264448, 264368, 264288, 21908785, 21908786, 21908787, 21908788, 21908789, 265020, 265021, 265022, 264682, 27486262, 27486284, 18108376, 20281152, 264558, 18108388, 87168518, 80432113, 22279000, 22279002, 264482
1756	82962614 (3511, 3512)	Novel Protein sim. GBank gi14432860 p AAC207081 - (AC008300) putative glucose-induced repressor protein (Arabidopsis thaliana)			22278997, 264259, 60432288, 29331827, 264908, 52844045, 285008, 264593, 285019, 264766, 21906766, 85274620, 18108385, 60432113, 264568, 264487
1757	85357380 (3513, 3514)	Novel Protein sim. GBank gi1541615 p CAB46858.1 - (AJ388557) zinc finger protein (Canis familiaris)	Contains protein domain (PF00096) - dna_mia_bind Zinc finger, C2H2 type		

1758	87612971 (3515, 3516)	Novel Protein sim. GBank gij381040lemj[CA416403] - (AL021487) predicted using GeneRinder [Caenorhabditis elegans]		UNCLASSIFIED	22278994, 22278999, 29331822, 29331824, 29331825, 33856970, 284508, 265006, 265007, 285009, 264591, 33657402, 33108954, 87168474, 284600, 285017, 285018, 21908788, 265020, 265021, 33657023, 33857109, 284629, 18108374, 35695855, 284632, 52844332, 22279002, 284583
1759	38894372 (3517, 3518)			UNCLASSIFIED	264759
1760	87328718 (3519, 3520)	Novel Protein sim. GBank gij5262748lemj[CA45688, 1] - (AJ133120) Proline rich synapse associated protein 2 [Rattus norvegicus]		UNCLASSIFIED	56182575, 60432049, 35696052, 264905, 264906, 264907, 264908, 264909, 265006, 265009, 264910, 60432229, 284592, 284595, 55812038, 284758, 264762, 18108351, 264764, 284786, 284788, 284789, 21906765, 55811957, 35695917, 264690, 264692, 264828, 284629, 55811578, 35696423, 264632, 264634, 284638, 284557, 284639, 60432113, 264585, 284488
1761	87408586 (3521, 3522)	Novel Protein sim. GBank gij127749spjP10568jMYSC_JCACA - MYOSIN IC HEAVY CHAIN	Contains protein domain (PF00818) - Guanine nucleotide exchange factor for Ras-like GTPases; N-terminal motif	UNCLASSIFIED	29331822, 284810, 284885, 284886
1762	95318867 (3523, 3524)	Novel Protein sim. GBank gij3169158 [AC004770] - BC289730_2 [Homo sapiens]	Contains protein domain (PF00173) - Heme-binding domain in cytochrome b5 and oxidoreductases	desaturase	22278999, 264259, 264905, 284907, 60170831, 285010, 285011, 285017, 284448, 21908765, 21908768, 21908767, 21908788, 265021, 284680, 33657109, 18108374, 284558, 60170394
1763	91224013 (3525, 3526)	Novel Protein sim. GBank gij4808026jph[AC004770] - (AF132858) suppressor of G2 allele of strp1 homolog [Homo sapiens]			56181686, 28331825, 35696052, 264905, 284906, 264909, 264763, 284682, 264769, 35695917, 265022, 33657023, 18108374, 35696423, 284834
1764	87757697 (3527, 3528)	Novel Protein sim. GBank gij1360669jph[CGHUTV - collagen alpha 1(V) chain precursor - human]	Contains protein domain (PF01410) - Fibrillar collagen C-terminal domain	collagen	22278998, 264239, 35696052, 29331828, 58182435, 285008, 285017, 265018, 264448, 264286, 21908768, 21908767, 29148627, 35695917, 264691, 33657023, 60432113, 22279002
1765	91230081 (3529, 3530)	Novel Protein sim. GBank gij486806jph[IS35503 - finger protein neutralized - fruit fly (Drosophila melanogaster)]		UNCLASSIFIED	264486, 18108394, 22278995, 22278997, 22278999, 284259, 264104, 264508, 284905, 284907, 284511, 285007, 285008, 284910, 285009, 21906754, 265010, 265017, 264603, 285018, 285019, 18108351, 284682, 264448, 264388, 264286, 284766, 18108358, 21908768, 21906767, 29148627, 29148629, 35695917, 265020, 265021, 264692, 264628, 264629, 18108374, 283978, 284636, 60373044, 22279000, 22279002

1766	85081201 (3531, 3532)	Novel Protein sim. GBank gi 2499087 sp C009337 UGGG - UDP- GLUCOSE:GLYCOPROTEIN GLUCOSYLTRANSFERASE PRECURSOR (DUGT)		glycoprotein	52645156, 67166559, 60170615, 33657023, 264693, 33657109, 27486261, 264555, 63373044
1767	87755998 (3533, 3534)	Novel Protein sim. GBank gi 176443 emb CAA18263.1 - (AL0222236) dJ1042K10.4 (novel protein) [Homo sapiens]		UNCLASSIFIED	265017, 265019, 264666, 264768, 265020, 264692
1768	80253216 (3535, 3536)			UNCLASSIFIED	264583
1769	67368968 (3537, 3538)	Novel Protein sim. GBank gi 115204 sp P00736 C1R_HUMAN - COMPLEMENT C1R COMPONENT PRECURSOR	Contains protein domain (PF00089) - Trypsin	UNCLASSIFIED	264583
1770	85413144 (3539, 3540)			UNCLASSIFIED	264583
1771	94233542 (3541, 3542)	Novel Protein sim. GBank gi 3914191 sp P56558 OQT1_RAT - UDP-N- ACETYLGLUCOSAMINE-PEPTIDE N- ACETYLGLUCOSAMINYLTRANSFERASE 110 KO SUBUNIT (O-GLCNAC TRANSFERASE P110 SUBUNIT) melanogaster]	Contains protein domain (PF00515) - TPR Domain	transferase	264756, 264600, 264369, 55611857, 265020, 63373044, 22279000
1772	67843510 (3543, 3544)	Novel Protein sim. GBank gi 4950442 gb AA04351.1 AF12136 - (AF121360) DNZDHH/CNEWT zinc finger protein 11 [Drosophila melanogaster]	Contains protein domain (PF01529) - DHHC zinc finger domain		22276996, 29331826, 33109854, 265018, 265019, 264764, 21906785, 265020, 265021, 264556
1773	94116824 (3545, 3546)	Novel Protein sim. GBank gi 397464 AF085693 - G protein-coupled receptor kinase-associated ADP ribosylation factor GTPase-activating protein [Rattus norvegicus]	Contains protein domain (PF01412) - Putative GTP-ase activating protein for Arf	potassium_channel	65274572, 56182575, 22278896, 35696286, 22278899, 264259, 29331824, 60424268, 29331825, 60432289, 35696052, 264106, 264509, 264906, 264907, 29331830, 264906, 52844045, 264511, 265006, 265007, 265008, 60170831, 60433436, 264756, 55611368, 67168558, 265017, 264604, 265019, 55611150, 264286, 56181562, 264669, 21906766, 21906767, 21906768, 21906769, 55611957, 265020, 265022, 52844150, 264691, 33657023, 264692, 264693, 60431528, 35696423, 35695855, 264636, 56182323, 18106367, 56526466, 22279000, 22279002, 264583, 264564, 264565, 264566, 264567
1774	94232573 (3547, 3548)	Novel Protein sim. GBank gi 2495689 sp C150341 032_HUMAN - HYPOTHETICAL PROTEIN KIAA0032	Contains protein domain (PF00415) - Regulator of chromosome condensation (RCC1)	UNCLASSIFIED	65274572, 56182575, 35696052, 55612038, 33109854, 21906754, 265017, 21906767, 21906768, 21906769, 265020, 264691, 264636, 56182323, 22279002

1775	95359330 (3548, 3550)	Novel Protein sim. GBank gll1489199db BAA094871 - (ID0928) The KIAA0138 gene product is novel. [Homo sapiens]		UNCLASSIFIED	65274572, 56182575, 56994075, 35686286, 22278987, 29331622, 29331828, 60432288, 28331628, 35880052, 29331830, 68712502, 264626, 58182435, 264511, 265007, 265008, 60170831, 80432229, 60433438, 55812038, 21908754, 85658542, 87166559, 264601, 265017, 265018, 285019, 264782, 264448, 264288, 264689, 21906765, 21906768, 21906767, 21906768, 265020, 265021, 265022, 264891, 33657023, 264892, 33657109, 27486261, 33857349, 18108370, 18108377, 35686423, 55811576, 35685855, 264632, 264634, 264836, 264639, 55162323, 83373044, 56526488, 87166518, 60432113, 22279000, 22279002, 264482, 264488, 264910
1776	94133756 (3551, 3552)	Novel Protein sim. GBank gll4588676 d BAA76857.1 - (AB023230) KIAA1013 protein [Homo sapiens]			264910
1777	87447171 (3553, 3554)	Novel Protein sim. GBank gll3219838 sp P87115 YDK9, SCHPO - HYPOTHETICAL 116.5 KD PROTEIN C20G8.09C IN CHROMOSOME 1	nuc1_recp1		56884075, 28331828, 265008, 87168474, 265017, 265018, 264761, 55811150, 264764, 56181582, 264689, 21908765, 21906768, 21906769, 35695917, 264690, 33657023, 35695763, 60431526, 35896423, 55811576, 35695855, 22279000, 22279002, 264564
1778	94851624 (3555, 3556)	Novel Protein sim. GBank gll3875648 emb CAA81454.1 - (Z68561) Similarity to Human rab13 protein (PIR Acc. No. A49647). Contains the ATP/GTP-binding site motif (PROSITE PS00017): cDNA EST EMBL: M69412 comes from this gene: cDNA EST YK212g9.3 comes from this gene: cDNA EST M212g9...	UNCLASSIFIED		29331828, 29331827, 35696052, 264512, 265007, 265009, 265017, 265018, 264782, 18108351, 264769, 21906765, 21908768, 21908767, 21908768, 21908768, 265020, 265021, 264691, 264693, 18108370, 264556, 83373044
1779	84133756 (3557, 3558)	Novel Protein sim. GBank gll4588676 d BAA76857.1 - (AB023230) KIAA1013 protein [Homo sapiens]	UNCLASSIFIED		293148827, 35686286, 29147620, 265008, 265007, 265008, 16108385, 65274727, 264482, 264369, 264769
1780	87023487 (3559, 3560)		UNCLASSIFIED		264107, 33657109, 56528466
1781	84047477 (3561, 3562)		Apidech		264508, 264906, 264639
1782	88094807 (3563, 3564)	Novel Protein sim. GBank gll729225 sp P41237 CTN_RAT - CORTEXIN	UNCLASSIFIED		264259, 28331622, 264508, 264805, 264906, 264907, 264908, 285007, 265009, 264910, 264591, 264758, 264764, 264286, 264788, 264769, 264635, 264636, 264637, 264639, 264563
1783	85717805 (3565, 3566)	Novel Protein sim. GBank gll2251543 d BAA214361 - (AB004538) protein arginine N-methyltransferase [Schizosaccharomyces pombe]	Interferon		264766

1764	95197093 (3567, 3568)	Novel Protein sim. GBank gi 1755049 (U55042) - myosin X [bos tauros]	Contains protein domain (PF00169) - PH domain	struct	35696288, 264259, 35698052, 264508, 264905, 264808, 264807, 66712502, 264908, 264909, 265007, 265008, 265009, 264910, 264591, 264594, 264757, 264758, 264759, 265010, 265011, 264601, 264602, 264804, 264805, 18108331, 264782, 264783, 264784, 264369, 264786, 264887, 264788, 264688, 21908768, 35695917, 264690, 264691, 264692, 264693, 264628, 18108374, 35696423, 264831, 264832, 264835, 264637, 264638, 264639, 18108385, 22278000, 22279002, 264565, 264568, 264468
1785	95357475 (3569, 3570)	Novel Protein sim. GBank gi K589552 db JAA76798.1 - (AB023171) KIAA0854 protein (Homo sapiens)		UNCLASSIFIED	65274572, 264259, 29331822, 29331824, 29331825, 29331827, 29331828, 264905, 264906, 264908, 66712502, 56182435, 264511, 265007, 60433356, 55811150, 264683, 264369, 264887, 52844229, 21908767, 52644150, 33657023, 65274620, 33857182, 85274791, 35695855, 264555, 65274727, 22279002
1786	85298485 (3571, 3572)	Novel Protein sim. GBank gi 17788 p P28770 CYA4_RAT ADENYLATE CYCLASE, TYPE IV (ATP PYROPHOSPHATE-LYASE) (ADENYLYL CYCLASE)		UNCLASSIFIED	264808, 35696423, 264638
1787	8734784 (3573, 3574)	Novel Protein sim. GBank gi 387175 emb CAA60338.1 - (Z50028) cDNA EST yk321h8.5 comes from this gene; cDNA EST EMBL:D88896 comes from this gene; cDNA EST yk395f9.5 comes from this gene [Caenorhabditis elegans]		UNCLASSIFIED	264488, 264905, 264908, 264909, 264585, 264764, 264766, 264892, 60431828, 264629, 264638, 264564, 264586
1788	91228778 (3575, 3576)	Novel Protein sim. GBank gi 2088669 (AF003130) - F55A12.9 gene product [Caenorhabditis elegans]		UNCLASSIFIED	264488, 63373044
1789	88084529 (3577, 3578)	Novel Protein sim. GBank gi 2088669 (AF003130) - F55A12.9 gene product [Caenorhabditis elegans]		UNCLASSIFIED	264488, 29331828, 264908, 18108351, 264288, 285021, 264555, 264638
1790	82489734 (3579, 3580)				35696052, 264905, 264906, 264907, 264908, 264909, 265008, 284910, 264758, 265011, 265019, 264764, 264766, 264769, 264826, 264835

1781	95187259 (3581, 3582)	Novel Protein sim. GBank gl1214321[db]BAAD0037 - (DB8733) membrane glycoprotein [Equine herpesvirus 1]	Contains protein domain (PF00047) - glycoprotein	264488, 264686, 264687, 264766, 16108334, 264769, 18106397, 264259, 264691, 264692, 33657023, 264693, 264508, 264805, 264806, 264628, 264907, 264829, 264908, 264909, 264510, 265008, 264511, 265008, 264830, 265009, 264631, 264910, 264832, 264834, 264635, 264555, 264636, 264592, 264637, 264593, 264638, 18106381, 264639, 264758, 265010, 265011, 264602, 22279000, 264604, 264760, 264584, 264681, 264762, 264565, 264763, 264683, 264566, 264764, 264286, 264684, 264567, 18108354, 18106391, 264685, 264766	
1792	87782690 (3583, 3584)	Novel Protein sim. GBank gl14337106[db]AAD18082 - (AF129756) BA-T4 [Homo sapiens]	Contains protein domain (PF01365) - G-patch domain	UNCLASSIFIED	22278997, 264259, 264508, 265007, 33657402, 87168559, 264369, 33657023, 35685855, 20281071, 264559, 18106387, 87168518
1793	95337877 (3585, 3586)	Novel Protein sim. GBank gl15579331[db]AAD45504, 1(AE14573) - (AF145732) endoplasmic reticulum alpha-mannosidase 1 [Homo sapiens]	Contains protein domain (PF01532) - Glycosyl hydrolase family 47	ATPase-associated	65274572, 22278895, 22278996, 22278997, 22278999, 264093, 264259, 28331824, 66714117, 60432289, 28331827, 29331828, 264103, 264105, 28331830, 285007, 264910, 265008, 60170831, 60433356, 21906754, 265010, 265017, 265019, 284681, 264682, 264286, 52644229, 21906765, 21906766, 21906767, 21906768, 21906769, 265020, 285021, 265022, 60170815, 52644150, 33657023, 33657109, 18108370, 18108374, 65274781, 20281071, 60432113, 22279000, 264482, 264564
1794	87759808 (3587, 3588)	Novel Protein sim. GBank gl14914604[db]CABA3677, 1 - (AL050389) hypothetical protein [Homo sapiens]	Contains protein domain (PF01798) - Putative snRNA binding domain	UNCLASSIFIED	18108394, 22278995, 22278999, 264259, 29331822, 29331824, 29331825, 29148486, 29146489, 264508, 264905, 52644045, 284112, 265006, 265008, 264910, 60433358, 264757, 55612038, 87168474, 285011, 285017, 16108351, 264763, 264448, 264683, 284369, 21906765, 21906768, 21906769, 21906770, 21906771, 21906772, 21906773, 21906774, 21906775, 21906776, 21906777, 21906778, 21906779, 21906780, 21906781, 21906782, 21906783, 21906784, 21906785, 21906786, 21906787, 21906788, 21906789, 21906790, 21906791, 21906792, 21906793, 21906794, 21906795, 21906796, 21906797, 21906798, 21906799, 21906800, 21906801, 21906802, 21906803, 21906804, 21906805, 21906806, 21906807, 21906808, 21906809, 21906810, 21906811, 21906812, 21906813, 21906814, 21906815, 21906816, 21906817, 21906818, 21906819, 21906820, 21906821, 21906822, 21906823, 21906824, 21906825, 21906826, 21906827, 21906828, 21906829, 21906830, 21906831, 21906832, 21906833, 21906834, 21906835, 21906836, 21906837, 21906838, 21906839, 21906840, 21906841, 21906842, 21906843, 21906844, 21906845, 21906846, 21906847, 21906848, 21906849, 21906850, 21906851, 21906852, 21906853, 21906854, 21906855, 21906856, 21906857, 21906858, 21906859, 21906860, 21906861, 21906862, 21906863, 21906864, 21906865, 21906866, 21906867, 21906868, 21906869, 21906870, 21906871, 21906872, 21906873, 21906874, 21906875, 21906876, 21906877, 21906878, 21906879, 21906880, 21906881, 21906882, 21906883, 21906884, 21906885, 21906886, 21906887, 21906888, 21906889, 21906890, 21906891, 21906892, 21906893, 21906894, 21906895, 21906896, 21906897, 21906898, 21906899, 21906900, 21906901, 21906902, 21906903, 21906904, 21906905, 21906906, 21906907, 21906908, 21906909, 21906910, 21906911, 21906912, 21906913, 21906914, 21906915, 21906916, 21906917, 21906918, 21906919, 21906920, 21906921, 21906922, 21906923, 21906924, 21906925, 21906926, 21906927, 21906928, 21906929, 21906930, 21906931, 21906932, 21906933, 21906934, 21906935, 21906936, 21906937, 21906938, 21906939, 21906940, 21906941, 21906942, 21906943, 21906944, 21906945, 21906946, 21906947, 21906948, 21906949, 21906950, 21906951, 21906952, 21906953, 21906954, 21906955, 21906956, 21906957, 21906958, 21906959, 21906960, 21906961, 21906962, 21906963, 21906964, 21906965, 21906966, 21906967, 21906968, 21906969, 21906970, 21906971, 21906972, 21906973, 21906974, 21906975, 21906976, 21906977, 21906978, 21906979, 21906980, 21906981, 21906982, 21906983, 21906984, 21906985, 21906986, 21906987, 21906988, 21906989, 21906990, 21906991, 21906992, 21906993, 21906994, 21906995, 21906996, 21906997, 21906998, 21906999, 21907000, 21907001, 21907002, 21907003, 21907004, 21907005, 21907006, 21907007, 21907008, 21907009, 21907010, 21907011, 21907012, 21907013, 21907014, 21907015, 21907016, 21907017, 21907018, 21907019, 21907020, 21907021, 21907022, 21907023, 21907024, 21907025, 21907026, 21907027, 21907028, 21907029, 21907030, 21907031, 21907032, 21907033, 21907034, 21907035, 21907036, 21907037, 21907038, 21907039, 21907040, 21907041, 21907042, 21907043, 21907044, 21907045, 21907046, 21907047, 21907048, 21907049, 21907050, 21907051, 21907052, 21907053, 21907054, 21907055, 21907056, 21907057, 21907058, 21907059, 21907060, 21907061, 21907062, 21907063, 21907064, 21907065, 21907066, 21907067, 21907068, 21907069, 21907070, 21907071, 21907072, 21907073, 21907074, 21907075, 21907076, 21907077, 21907078, 21907079, 21907080, 21907081, 21907082, 21907083, 21907084, 21907085, 21907086, 21907087, 21907088, 21907089, 21907090, 21907091, 21907092, 21907093, 21907094, 21907095, 21907096, 21907097, 21907098, 21907099, 21907100, 21907101, 21907102, 21907103, 21907104, 21907105, 21907106, 21907107, 21907108, 21907109, 21907110, 21907111, 21907112, 21907113, 21907114, 21907115, 21907116, 21907117, 21907118, 21907119, 21907120, 21907121, 21907122, 21907123, 21907124, 21907125, 21907126, 21907127, 21907128, 21907129, 21907130, 21907131, 21907132, 21907133, 21907134, 21907135, 21907136, 21907137, 21907138, 21907139, 21907140, 21907141, 21907142, 21907143, 21907144, 21907145, 21907146, 21907147, 21907148, 21907149, 21907150, 21907151, 21907152, 21907153, 21907154, 21907155, 21907156, 21907157, 21907158, 21907159, 21907160, 21907161, 21907162, 21907163, 21907164, 21907165, 21907166, 21907167, 21907168, 21907169, 21907170, 21907171, 21907172, 21907173, 21907174, 21907175, 21907176, 21907177, 21907178, 21907179, 21907180, 21907181, 21907182, 21907183, 21907184, 21907185, 21907186, 21907187, 21907188, 21907189, 21907190, 21907191, 21907192, 21907193, 21907194, 21907195, 21907196, 21907197, 21907198, 21907199, 21907200, 21907201, 21907202, 21907203, 21907204, 21907205, 21907206, 21907207, 21907208, 21907209, 21907210, 21907211, 21907212, 21907213, 21907214, 21907215, 21907216, 21907217, 21907218, 21907219, 21907220, 21907221, 21907222, 21907223, 21907224, 21907225, 21907226, 21907227, 21907228, 21907229, 21907230, 21907231, 21907232, 21907233, 21907234, 21907235, 21907236, 21907237, 21907238, 21907239, 21907240, 21907241, 21907242, 21907243, 21907244, 21907245, 21907246, 21907247, 21907248, 21907249, 21907250, 21907251, 21907252, 21907253, 21907254, 21907255, 21907256, 21907257, 21907258, 21907259, 21907260, 21907261, 21907262, 21907263, 21907264, 21907265, 21907266, 21907267, 21907268, 21907269, 21907270, 21907271, 21907272, 21907273, 21907274, 21907275, 21907276, 21907277, 21907278, 21907279, 21907280, 21907281, 21907282, 21907283, 21907284, 21907285, 21907286, 21907287, 21907288, 21907289, 21907290, 21907291, 21907292, 21907293, 21907294, 21907295, 21907296, 21907297, 21907298, 21907299, 21907300, 21907301, 21907302, 21907303, 21907304, 21907305, 21907306, 21907307, 21907308, 21907309, 21907310, 21907311, 21907312, 21907313, 21907314, 21907315, 21907316, 21907317, 21907318, 21907319, 21907320, 21907321, 21907322, 21907323, 21907324, 21907325, 21907326, 21907327, 21907328, 21907329, 21907330, 21907331, 21907332, 21907333, 21907334, 21907335, 21907336, 21907337, 21907338, 21907339, 21907340, 21907341, 21907342, 21907343, 21907344, 21907345, 21907346, 21907347, 21907348, 21907349, 21907350, 21907351, 21907352, 21907353, 21907354, 21907355, 21907356, 21907357, 21907358, 21907359, 21907360, 21907361, 21907362, 21907363, 21907364, 21907365, 21907366, 21907367, 21907368, 21907369, 21907370, 21907371, 21907372, 21907373, 21907374, 21907375, 21907376, 21907377, 21907378, 21907379, 21907380, 21907381, 21907382, 21907383, 21907384, 21907385, 21907386, 21907387, 21907388, 21907389, 21907390, 21907391, 21907392, 21907393, 21907394, 21907395, 21907396, 21907397, 21907398, 21907399, 21907400, 21907401, 21907402, 21907403, 21907404, 21907405, 21907406, 21907407, 21907408, 21907409, 21907410, 21907411, 21907412, 21907413, 21907414, 21907415, 21907416, 21907417, 21907418, 21907419, 21907420, 21907421, 21907422, 21907423, 21907424, 21907425, 21907426, 21907427, 21907428, 21907429, 21907430, 21907431, 21907432, 21907433, 21907434, 21907435, 21907436, 21907437, 21907438, 21907439, 21907440, 21907441, 21907442, 21907443, 21907444, 21907445, 21907446, 21907447, 21907448, 21907449, 21907450, 21907451, 21907452, 21907453, 21907454, 21907455, 21907456, 21907457, 21907458, 21907459, 21907460, 21907461, 21907462, 21907463, 21907464, 21907465, 21907466, 21907467, 21907468, 21907469, 21907470, 21907471, 21907472, 21907473, 21907474, 21907475, 21907476, 21907477, 21907478, 21907479, 21907480, 21907481, 21907482, 21907483, 21907484, 21907485, 21907486, 21907487, 21907488, 21907489, 21907490, 21907491, 21907492, 21907493, 21907494, 21907495, 21907496, 21907497, 21907498, 21907499, 21907500, 21907501, 21907502, 21907503, 21907504, 21907505, 21907506, 21907507, 21907508, 21907509, 21907510, 21907511, 21907512, 21907513, 21907514, 21907515, 21907516, 21907517, 21907518, 21907519, 21907520, 21907521, 21907522, 21907523, 21907524, 21907525, 21907526, 21907527, 21907528, 21907529, 21907530, 21907531, 21907532, 21907533, 21907534, 21907535, 21907536, 21907537, 21907538, 21907539, 21907540, 21907541, 21907542, 21907543, 21907544, 21907545, 21907546, 21907547, 21907548, 21907549, 21907550, 21907551, 21907552, 21907553, 21907554, 21907555, 21907556, 21907557, 21907558, 21907559, 21907560, 21907561, 21907562, 21907563, 21907564, 21907565, 21907566, 21907567, 21907568, 21907569, 21907570, 21907571, 21907572, 21907573, 21907574, 21907575, 21907576, 21907577, 21907578, 21907579, 21907580, 21907581, 21907582, 21907583, 21907584, 21907585, 21907586, 21907587, 21907588, 21907589, 21907590, 21907591, 21907592, 21907593, 21907594, 21907595, 21907596, 21907597, 21907598, 21907599, 21907600, 21907601, 21907602, 21907603, 21907604, 21907605, 21907606, 21907607, 21907608, 21907609, 21907610, 21907611, 21907612, 21907613, 21907614, 21907615, 21907616, 21907617, 21907618, 21907619, 21907620, 21907621, 21907622, 21907623, 21907624, 21907625, 21907626, 21907627, 21907628, 21907629, 21907630, 21907631, 21907632, 21907633, 21907634, 21907635, 21907636, 21907637, 21907638, 21907639, 21907640, 21907641, 21907642, 21907643, 21907644, 21907645, 21907646, 21907647, 21907648, 21907649, 21907650, 21907651, 21907652, 21907653, 21907654, 21907655, 21907656, 21907657, 21907658, 21907659, 21907660, 21907661, 21907662, 21907663, 21907664, 21907665, 21907666, 21907667, 21907668, 21907669, 21907670, 21907671, 21907672, 21907673, 21907674, 21907675, 21907676, 21907677, 21907678, 21907679, 21907680, 21907681, 21907682, 21907683, 21907684, 21907685, 21907686, 21907687, 21907688, 21907689, 21907690, 21907691, 21907692, 21907693, 21907694, 21907695, 21907696, 21907697, 21907698, 21907699, 21907700, 21907701, 21907702, 21907703, 21907704, 21907705, 21907706, 21907707, 21907708, 21907709, 21907710, 21907711, 21907712, 21907713, 21907714, 21907715, 21907716, 21907717, 21907718, 21907719, 21907720, 21907721, 21907722, 21907723, 21907724, 21907725, 21907726, 21907727, 21907728, 21907729, 21907730, 21907731, 21907732, 21907733, 21907734, 21907735, 21907736, 21907737, 21907738, 21907739, 21907740, 21907741, 21907742, 21907743, 21907744, 21907745, 21907746, 21907747, 21907748, 21907749, 21907750, 21907751, 21907752, 21907753, 21907754, 21907755, 21907756, 21907757, 21907758, 21907759, 21907760, 21907761, 21907762, 21907763, 21907764, 21907765, 21907766, 21907767, 21907768, 21907769, 21907770, 21907771, 21907772, 21907773, 21907774, 21907775, 21907776, 21907777, 21907778, 21907779, 21907780, 21907781, 21907782, 21907783, 21907784, 21907785, 21907786, 21907787, 21907788, 21907789, 21907790, 21907791, 21907792, 21907793, 21907794, 21907795, 21907796, 21907797, 21907798, 21907799, 21907800, 21907801, 21907802, 21907803, 21907804, 21907805, 21907806, 21907807, 21907808, 21907809, 21907810, 21907811, 21907812, 21907813, 21907814, 21907815, 21907816, 21907817, 21907818, 21907819, 21907820, 21907821,

1797	91223219 (3593, 3594)	Novel Protein sim. GBank g11842111 (U87566) - decoy [Arabidopsis thaliana]		ribosomalprot	22278996, 22278997, 22278998, 22278999, 29331822, 284910, 6017031, 21906754, 52844229, 21906765, 21906768, 21908769, 35695917, 265022, 52644150, 264691, 33857023, 263967, 33657109, 22279000
1798	91221276 (3595, 3596)	Novel Protein sim. GBank g12832906dubj1AA24608.11 - (D69340) dipeptidyl peptidase III [Rattus norvegicus]		peptidase	22278994, 56994075, 22278997, 22278998, 22278999, 264259, 28331828, 60432268, 28331828, 33856970, 265008, 60432229, 264757, 60433438, 21906754, 33657084, 87168559, 265017, 18108351, 264882, 264446, 264268, 21906765, 21906766, 21908767, 21906768, 21908769, 35695917, 265020, 265021, 33857023, 33857182, 27486261, 27486265, 33657349, 263973, 18108374, 55811576, 35895855, 18106385, 87168516, 22279000, 284486
1799	86321713 (3597, 3598)	Novel Protein sim. GBank g15689541dbj1BAA83054.11 - (AB028025) KIAA1102 protein [Homo sapiens]		eph	284908, 21906754, 21906767, 21908769, 265020, 33657023, 264692, 264693, 264404, 22279000
1800	87080118 (3599, 3600)	Novel Protein sim. GBank g14680679dbj1AAD27728.11AF13295 - (AF132954) CGI-20 protein [Homo sapiens]		UNCLASSIFIED	264691, 264558, 264566 52844045, 265007, 264832
1801	95060723 (3601, 3602)	Novel Protein sim. GBank g1134920sp1P21897SSGP_VOLCA - SULFATED SURFACE GLYCOPROTEIN 185 (SSG 185)			35696286, 66714117, 264508, 264509, 56182435, 264512, 18108351, 264688, 55811957, 264692, 55811578, 35695855, 264486
1802	87771012 (3603, 3604)	Novel Protein sim. GBank g14680679dbj1AAD27728.11AF13295 - (AF132954) CGI-20 protein [Homo sapiens]			264688, 264468, 264687, 264488, 284788, 264769, 264689, 21908769, 35696288, 35895917, 264259, 264691, 264692, 264693, 20281099, 18108384, 35896052, 264508, 264508, 264905, 264906, 18108370, 264628, 264907, 66712502, 264908, 264909, 18108374, 18108376, 35896423, 35695855, 264510, 264511, 265006, 265007, 264512, 285008, 264910, 264631, 264632, 264634, 264635, 264591, 264636, 264637, 264592, 264636, 264593, 264638, 264594, 63373044, 264758, 264598, 18108385, 18108387, 265011, 264780, 264593, 18108351, 264782, 264584, 264448, 264585, 284783, 264883, 284764, 264588, 264288, 264486, 264587, 264785, 264766
1803	95060725 (3605, 3606)	Novel Protein sim. GBank g14680679dbj1AAD27728.11AF13295 - (AF132954) CGI-20 protein [Homo sapiens]			

1804	87770203 (3607, 3608)	Novel Protein sim. GBank gij3878914fembjCAA9538.11 - (Z74043) predicted using GeneFinder; cDNA EST EMBL: C13850 comes from this gene; cDNA EST EMBL: C11575 comes from this gene; cDNA EST yk3434.5 comes from this gene [Caenorhabditis elegans]			52646365, 22278997, 22278999, 264905, 264908, 264909, 264910, 21908754, 264766, 21908785, 21908788, 35695917, 265020, 265022, 264691, 264637, 264639, 22279000, 264564, 264566
1805	95330375 (3609, 3610)	Novel Protein sim. GBank gij5453644refjNP_006461.1PEBBP - estrogen-responsive B box protein			29331824, 29331825, 29331828, 29331827, 29331828, 87168559, 264288, 264687, 52844229, 35698423, 264638, 60432113
1806	94133782 (3611, 3612)	Novel Protein sim. GBank gij4589876dbjBAA78857.11 - (AB023230) KIAA1013 protein [Homo sapiens]		struct	264094, 264105, 264908, 35888423, 285006, 265007, 265008, 264555, 264592, 265011, 265018, 264369
1807	88943032 (3613, 3614)				29331824, 264908, 264910, 33657023, 263976
1808	87642711 (3615, 3616)	Novel Protein sim. GBank gij4864079fembjCAB43235.11 - (AL050008) hypothetical protein [Homo sapiens]		UNCLASSIFIED	264468, 35698286, 86714117, 35698052, 66712502, 264592, 60433438, 52844296, 265010, 264683, 264369, 264689, 55811937, 35895917, 33857109, 35695783, 55810764, 18108379, 35886423, 35895855, 58182323, 284583, 264584, 264487
1809	95321488 (3617, 3618)	Novel Protein sim. GBank gij1918927 (U87965) - putative G protein [Mus musculus]		UNCLASSIFIED	264594, 55811150, 264886, 29148629, 29148784, 264690, 264628, 18108374, 264556, 264557, 264558
1810	88066318 (3619, 3620)	Novel Protein sim. GBank gij135294450p47179YJ9P - YEAST 118.4 KD PROTEIN IN BATZ-DALS INTERGENIC REGION PRECURSOR		UNCLASSIFIED	264488, 35698052, 264905, 264906, 264907, 264908, 264909, 264511, 265008, 264910, 264592, 264593, 264594, 33657402, 284757, 264595, 264758, 264598, 264759, 264600, 264601, 284782, 264683, 264784, 264288, 264684, 264788, 264787, 264808, 264788, 264887, 284769, 264689, 265021, 264690, 264691, 264693, 264628, 264629, 18108374, 264630, 264631, 264632, 264634, 264635, 264636, 284837, 264838, 264639, 264563, 264566, 284486, 264567
1811	88086272 (3621, 3622)	Novel Protein sim. GBank gij12134984pji137275 - death-associated protein kinase (EC 2.7.1.-) - human	Contains protein domain (PF00023) - kinase Ank repeat	kinase	264488, 264259, 264508, 264509, 264905, 264908, 284907, 58182435, 264511, 264512, 284910, 264756, 265011, 264600, 264604, 18108354, 264768, 264888, 284768, 264534, 60170615, 33657023, 264829, 264831, 264639, 284563, 284482, 284483
1812	79245772 (3623, 3624)				29331822, 29331824, 265019, 18108351, 21908769

1813	88090972 (3625, 3626)	Novel Protein sim. GBank gij5051636gijAAD38328.1(AF073727) EH domain-binding mitotic phosphoprotein [Homo sapiens]	Contains protein domain (PF01417) - ENTH domain	glucoamylase	56182575, 264229, 29331824, 66714117, 29331828, 35695052, 264509, 264905, 264908, 264907, 264908, 66712502, 264809, 265007, 264910, 264591, 264593, 55612038, 265011, 265018, 264780, 264662, 264764, 264883, 264359, 264786, 264788, 264789, 21906766, 21906786, 264691, 264693, 18108374, 35695855, 264634, 264635, 264637, 264639, 264559, 22279000, 22279002, 264568
1614	86176047 (3627, 3628)	Novel Protein sim. GBank gij3643808 (AC005395) - hypothetical protein [Arabidopsis thaliana]		UNCLASSIFIED	264488, 35696286, 22278988, 264092, 264094, 284229, 29331822, 29331824, 29331825, 60432289, 29331826, 29331827, 29331828, 264105, 264107, 52644045, 56182435, 265009, 60432229, 60433356, 87166474, 87166559, 264389, 264288, 21906765, 35695917, 265021, 285022, 33657023, 33657109, 18106374, 35696423, 264638, 58528486, 264462
1615	65286473 (3629, 3630)	Novel Protein sim. GBank gij117786gijP26770CYAD_RAT ADENYLATE CYCLASE, TYPE IV (ATP PYROPHOSPHATE-LYASE) (ADENYLYL CYCLASE)		struct	22278999, 264508, 264509, 264907, 264908, 264910, 265011, 264780, 264786, 264634, 264836
1818	83738645 (3631, 3632)	Novel Protein sim. GBank gij1176623gijP41848Y098_CAEEL - HYPOTHETICAL 93.9 KD PROTEIN T20812.6 IN CHROMOSOME III		UNCLASSIFIED	18108394, 18108397, 264509, 264907, 264908, 264909, 265009, 264591, 265011, 285017, 284687, 264689, 285022, 264691, 18106362, 18106366, 18106370, 18106374, 18106378, 264635, 264557, 264584, 264587
1817	88095286 (3633, 3634)	Novel Protein sim. GBank gij3766377gijmbjCAA214291 - (AL031907) hypothetical protein [Schizosaccharomyces pombe]	Contains protein domain (PF00400) - WD domain, G-beta repeat	histone	264488, 22278997, 22278999, 60432049, 29331822, 29331824, 60432288, 52644045, 60170831, 265017, 265016, 265019, 18108351, 264862, 52644228, 21906765, 21906767, 21906766, 52644150, 33657023, 33657109, 27486282, 18106370, 18106374, 60170394, 56182323, 22279002
1818	85806775 (3635, 3636)	Novel Protein sim. GBank gij3679121gijmbjCAA943701 - (Z70310) predicted using GeneFinder. Similarity to Mouse ankyrin (PIR Acc. No. S37771); cDNA EST EMBL: T01923 comes from this gene; cDNA EST EMBL: D32335 comes from this gene; cDNA EST EMBL: D32723 comes from this gene; cDNA EST E...	Contains protein domain (PF00023) - Ank repeat	transcriptfactor	35686286, 60433356, 264758, 264389, 264688, 21906769, 264693, 264632
1819	67758572 (3637, 3638)	Novel Protein sim. GBank gij5031865gijNP_005771.1pLHFP - lipoma HMGIC fusion partner		UNCLASSIFIED	22278995, 22278997, 22278999, 264259, 29331826, 264908, 285007, 265008, 265009, 60432228, 265017, 265018, 265019, 264448, 264286, 21906788, 21906789, 265020, 18108361, 18108364, 22278900, 22279002, 264567

1820	87789455 (3638, 3640)					264905, 264907, 264594
1821	80431510 (3641, 3642)					264807, 264768, 263976
1822	91221523 (3643, 3644)	Novel Protein sim. GBank g14864130[emb]CAB43272.11- (AL050101) hypothetical protein [Homo sapiens]				22278995, 56994075, 22278996, 22278997, 22278998, 264259, 29331824, 29331825, 29331826, 35696052, 29331826, 264908, 29331830, 60170831, 264591, 264593, 60433358, 264596, 265017, 265019, 18108351, 264763, 264883, 21906765, 21906767, 21906768, 21906769, 35695917, 265020, 265021, 33657023, 18108384, 16108370, 35693855, 22279000, 22279002 264488, 264259, 264511, 264268, 264768, 264693, 35696423, 264634, 18108385, 264488
1823	85522330 (3645, 3646)				UNCLASSIFIED	264807, 264808, 264809, 264511, 264631, 264634, 264835, 264637, 264638, 264639, 264758, 264566
1824	86612025 (3647, 3648)	Novel Protein sim. GBank g1477072[pf]A48018 - mucin 7 precursor, salivary - human	Contains protein domain (PF00047) - immunoglobulin domain		UNCLASSIFIED	60432049, 264910, 264487
1825	87430125 (3649, 3650)	Novel Protein sim. GBank g13036903[emb]CAA18493 - (AL022373) hypothetical protein [Arabidopsis thaliana]			UNCLASSIFIED	52844507, 52645156, 52846842, 22278994, 22278996, 56994075, 264259, 60432049, 22278998, 56994075, 264259, 60432049, 52845080, 35696052, 66712502, 52644045, 265008, 265009, 60432229, 80433358, 60433438, 52646317, 52644296, 265011, 87168559, 264448, 264268, 264369, 264688, 52644229, 264689, 21906765, 21906768, 285020, 60170815, 52644150, 33657023, 27488262, 27488264, 27488265, 35695763, 35696423, 35695855, 83373044, 87168516, 264404, 22279002
1826	91723612 (3651, 3652)	Novel Protein sim. GBank g14680885[gb]AACD27732.11AF13295 - (AF132957) CGI-23 protein [Homo sapiens]			ATPase-associated	264758
1827	81647212 (3653, 3654)					264758

1828	85074017 (3855, 3858)	Novel Protein sim. GBank gll4503571[en]NP_001419, 1pEN01 - endase 1, (alpha)	Contains protein domain (PF00113) - Endase	oncogene	264488, 52646842, 58182575, 22278996, 3588928, 22278997, 22278998, 284091, 264093, 60432048, 264259, 29331822, 29331824, 88714117, 29331825, 80432289, 29331828, 29331827, 29331828, 264105, 264508, 264807, 68712502, 52844045, 58182435, 265006, 264511, 284512, 265007, 265008, 265009, 80170831, 60432229, 264593, 80433358, 80433438, 264758, 33109954, 21908754, 87188474, 265010, 265011, 87188558, 265017, 265019, 264781, 264782, 264448, 264784, 264883, 264288, 264369, 18108355, 264769, 264689, 18108358, 264888, 264769, 264689, 21906788, 21906789, 35895917, 285021, 60170615, 33657023, 33657349, 283972, 55811578, 35895855, 284635, 264555, 264558, 264638, 264557, 87188518, 22279000, 22279002, 264563, 264482, 264585, 264484, 284587, 284508, 264834, 284509, 284482, 29331827, 264908, 285009, 264910
1829	80197720 (3857, 3858)				264508, 264834, 284509, 284482, 29331827, 264908, 285009, 264910
1830	94312942 (3859, 3860)	Novel Protein sim. GBank gll2245532 (U58755) - ORF 73, contains large complex repeat CR 73 [Kaposi's sarcoma-associated herpesvirus]		nuclease	52645156, 22278994, 22278995, 35896288, 22278996, 22278997, 22278998, 22278999, 29331822, 29331825, 35896052, 52646317, 52644296, 87188559, 265018, 21906785, 21908788, 21908787, 21906788, 21906789, 35895917, 265021, 33657023, 52645129, 33657109, 33657182, 27488281, 27488282, 35895763, 263974, 35896423, 35895855, 52644332
1831	84138063 (3861, 3862)			UNCLASSIFIED	29331824, 35896052, 29331830, 284595, 264758, 265010, 265018, 285022, 264883, 65274791
1832	84521663 (3863, 3864)	Novel Protein sim. GBank gll1330345 (U58755) - coded for by C. elegans cDNA yk34b1.5; coded for by C. elegans cDNA yk13h10.5; coded for by C. elegans cDNA yk46d5.5; coded for by C. elegans cDNA yk43c2.5; coded for by C. elegans cDNA yk46e8....			264602

1833	85314184 (3665, 3666)	Novel Protein sim. GBank gll15174413jfeINP_006028.1 PCDC4 - CDC42-binding protein kinase beta (DMPK-like)	Contains protein domain (PF00069) - Kinase Eukaryotic protein kinase domain	22278994, 22278997, 22278998, 22278999, 264259, 28331822, 29331824, 60432288, 29331827, 35896052, 29146489, 264508, 284508, 264808, 264807, 60712502, 294908, 52844045, 264909, 284512, 265008, 264581, 264593, 60433358, 21908754, 33657084, 265011, 285017, 264604, 285018, 285019, 264681, 18108351, 284883, 264288, 264885, 264768, 284687, 21906765, 21906766, 21906767, 21906768, 21906769, 29148829, 285020, 285021, 284690, 284892, 33657023, 65274791, 284834, 284635, 284556, 284557, 264558, 284559, 18108385, 56526488, 87168518, 60432113, 22278000, 22278002, 264563
1834	80562790 (3667, 3668)			264259, 264807, 264889, 22279000, 22279002
1835	94135718 (3669, 3670)		UNCLASSIFIED	22278998, 28331822, 29331828, 87168474, 284803, 21908768, 283978, 35895855, 83373044
1836	87348450 (3671, 3672)	Novel Protein sim. GBank gll429288 NIP_004288.1 PUCP4 - uncoupling protein 4	Contains protein domain (PF00153) - transport Mitochondrial carrier proteins	28331825, 264908, 285018, 264784, 284886, 21908765, 264635
1837	94234297 (3673, 3674)	Novel Protein sim. GBank gll334400 pIC24574 UBE, DROME - UBIQUITIN CARBOXYL-TERMINAL HYDROLASE 64E (UBIQUITIN THIOLESTERASE 64E) (UBIQUITIN-SPECIFIC PROCESSING PROTEASE 64E) (DEUBIQUITINATING ENZYME 64E)	Contains protein domain (PF00443) - ubiquitin Ubiquitin carboxyl-terminal hydrolase family 2	22278895, 29146498, 265006, 265008, 265009, 285010, 264883, 21908765, 29148627, 29148829, 285020, 265021, 285022, 65274820, 18108370, 18108374, 284558, 18108385
1838	94324369 (3675, 3676)	Novel Protein sim. GBank gll1382599 pI A58154 - Axl substrate ena (enabled) - fruit fly (Drosophila melanogaster)	Contains protein domain (PF00568) - WH1 domain	28331822, 265017, 264760, 285020, 83373044
1839	87456508 (3677, 3678)	Novel Protein sim. GBank gll2117310 emb CA808118.11 - (Z95620) hypothetical protein [Schizosaccharomyces pombe]	UNCLASSIFIED	60433438, 264601, 21906765, 21906766, 265021, 33657109, 284556
1840	87391708 (3679, 3680)	Novel Protein sim. GBank gll127560 p P23249 MV10 MOUSE - PROTEIN MOV-10	UNCLASSIFIED	284693
1841	85818445 (3681, 3682)	Novel Protein sim. GBank gll4572484 p AUC23834.1 AF12385 - (AF12385) FEZ1 [Homo sapiens]		58182575, 29331824, 28331828, 60433356, 264764, 264288, 33657023, 283987, 18108370, 18108374, 284831, 284555, 284556, 284639

1842	80982845 (3683, 3684)	Novel Protein sim. GBank g11326286 (U58728) - C54H2.1 gene product [<i>Caenorhabditis elegans</i>]		UNCLASSIFIED	652745172, 29331822, 29331824, 29331825, 68714117, 29331826, 29331827, 264907, 264908, 52644045, 56182435, 264510, 265008, 265007, 265009, 264910, 60433356, 264757, 60433438, 55812036, 265017, 18108354, 264688, 264789, 33657023, 264693, 18106384, 33857109, 18108368, 264828, 55810764, 56182323, 18106384, 264563, 264564
1843	85282892 (3685, 3686)			UNCLASSIFIED	264486, 56182435, 264769, 29331828, 29331826, 264511, 265006, 265007, 264910, 264631, 264509, 264680, 264636, 264584, 284691, 60432229, 60432049, 264259, 264629, 33657023, 264488, 264909, 284567, 264595, 264766
1844	67444764 (3687, 3688)	Novel Protein sim. GBank g124988875p1009232YQ22, CAEL - HYPOTHETICAL 32.0 KD PROTEIN C09F5.2 IN CHROMOSOME III		UNCLASSIFIED	264906, 265022, 33657023, 87188518, 22279002
1845	95096673 (3689, 3690)	Novel Protein sim. GBank g1175494pp1008819YAC5, SCHPO - HYPOTHETICAL 45.0 KD PROTEIN C16C8.05 IN CHROMOSOME I	Contains protein domain (PF00628) - PHD-finger	transcriptfactor	264259, 29331824, 264907, 264908, 66712502, 264510, 265007, 265006, 55812038, 265018, 21808785, 52844150, 33657109, 264555, 264558, 264557, 56182323, 18106382, 83373044, 18108385, 264564
1846	84267872 (3691, 3692)	Novel Protein sim. GBank g138610801emb1CA217391 - (A1.032657) similar to EGF-like domain, CDNA EST yk299a12.3 comes from this gene; CDNA EST EMBL.D35388 comes from this gene; CDNA EST yk33116.5 comes from this gene; CDNA EST yk299a12.5 comes from this gene; CDNA EST yk46718.5...	Contains protein domain (PF00008) - EGF-like domain	UNCLASSIFIED	264905, 264908
1847	87621487 (3693, 3694)	Novel Protein sim. GBank g15059323jg1AD38867.1AF15152 - (AF151522) helix end enhancer of split related-1 [Homo sapiens]	Contains protein domain (PF00010) - Helix-loop-helix DNA-binding domain	transcriptfactor	22278997, 264259, 29331824, 264909, 18108351, 263974, 22279002
1848	86789380 (3695, 3696)	Novel Protein sim. GBank g157016541emb1CAB52191.11 - (AJ245417) G5b protein [Homo sapiens]		tm7	29331825, 29331826, 29331827, 265017, 264983, 264288, 264768, 264768, 21806767, 21908766, 264692, 22279002
1849	84287874 (3697, 3698)	Novel Protein sim. GBank g145036651ref1NP_001989.1PFBLN - fibulin 2 precursor	Contains protein domain (PF00008) - EGF-like domain	ATPase_associated	56182575, 265018
1850	88689650 (3699, 3700)	Novel Protein sim. GBank g145685621db1BA476813.11 - (AB023166) KIAA0969 protein [Homo sapiens]	Contains protein domain (PF00169) - PH domain	strud	60432049, 264908
1851	85419789 (3701, 3702)	Novel Protein sim. GBank g1220837db1BA4014771 - (D10827) zinc finger protein [Mus musculus]	Contains protein domain (PF00086) - Zinc finger, C2H2 type	dna_ma_bind	29331824, 35696052, 264910, 60433438, 264688, 35695917, 265020, 52644150, 86274620, 52644332

1852	95413170 (3703, 3704)	Novel Protein sim. GBank glt5774629[eflNP_006090.1]pPIAS - protein inhibitor of activated STAT3		UNCLASSIFIED	56182575, 35686286, 22278998, 22278997, 22278999, 284490, 80432049, 264259, 29331822, 29331824, 29331625, 29331626, 60432289, 29331627, 35696052, 52844045, 285007, 264910, 60432229, 60433358, 60433438, 55812038, 65274444, 265018, 285019, 18108351, 284448, 284688, 264687, 21908765, 21908767, 21908769, 265021, 265022, 52644150, 264693, 33657109, 18108370, 18108374, 55811578, 35695855, 56182323, 80432113, 22279002, 284383
1853	91222287 (3705, 3706)	Novel Protein sim. GBank glt654085[emb]CAA583371 - (X63413) U88 [Human herpesvirus 6]		UNCLASSIFIED	284687, 264768, 52644507, 264769, 21908765, 21906787, 21906768, 22278995, 56994075, 22278999, 52644150, 264259, 264692, 29331822, 29331824, 52645129, 29331827, 33656970, 33657349, 35695783, 264508, 264906, 264628, 264907, 264629, 264909, 35699423, 35695855, 284510, 265008, 284511, 264512, 284630, 265009, 264631, 284910, 284634, 284835, 284637, 264593, 264838, 264639, 33657402, 18108365, 52646317, 52844298, 87188516, 67168558, 264602, 265017, 22279000, 265018, 264760, 264782, 284882, 264448, 264784, 264684, 284587, 284288, 264389, 264786
1854	88036152 (3707, 3708)	Novel Protein sim. GBank glt2072864 (U83569) - putative p150 [Homo sapiens]		nuclease	284592
1855	91221459 (3709, 3710)	Novel Protein sim. GBank glt4539520[emb]CAB39994.1 - (AL035424) DA22D12.1 (novel protein similar to Drosophila Kelch (RING Canal protein, KEL) and a heterogeneous set of other types of proteins) [Homo sapiens]	Contains protein domain (PF01344) - Kelch motif	nuc. recpt	18108392, 52646365, 65274572, 56182575, 22278994, 56994075, 22278999, 35696288, 22278998, 284259, 60432049, 264608, 284510, 285007, 265008, 285009, 264595, 21906754, 87168474, 285011, 87168559, 264681, 264288, 264768, 18108359, 21906784, 21906768, 28148627, 285020, 265021, 265022, 52644150, 33657023, 33657109, 18108372, 18108374, 18108376, 35696423, 264831, 284638, 18108361, 264482
1856	94231671 (3711, 3712)	Novel Protein sim. GBank glt395497[emb]CAA06945 - (AJ008276) acetylglucosaminyltransferase-like protein [Mus musculus]		UNCLASSIFIED	56994075, 264259, 29331826, 264511, 284910, 264758, 264693, 264637, 18108361, 63373044
1857	94324455 (3713, 3714)	Novel Protein sim. GBank glt4322670[gb]AAD161201 - (AF094508) dentin phosphophoryn [Homo sapiens]		ATPase-associated	22278999, 264258, 284608, 60170831, 264448, 264686, 265020, 285022, 33657109, 60170394, 83373044

1856	87828311 (3715, 3716)	Novel Protein sim. GBank gij4981903jpbIAAD36415_1(AE001789)	Contains protein domain (PF00312) - ribosomalprot	284757
1859	84407464 (3717, 3718)	Novel Protein sim. GBank gij4240317jdbjBAAT4937_11-(AB020721) KIAA0914 protein [Homo sapiens]		22278998, 28331824, 265007, 33109954, 265019, 284369, 21908768, 29148784, 27486261, 52644332, 22279002
1860	17929308 (3719, 3720)	Novel Protein sim. GBank gij4009522 (AF099731) - connexin 31.1 [Homo sapiens]		265019
1861	88086370 (3721, 3722)	Novel Protein sim. GBank gij2143637jpirjil84505 - calcium-dependent actin-binding protein - lat	Contains protein domain (PF00285) - Citrate synthase	264887, 264259, 29331822, 29331824, 29331825, 265007, 285009, 284591, 33109954, 285010, 265019, 284389, 284288, 264686, 284691, 264693, 27486284, 18108370, 18108374, 283977, 55811578, 56182323, 284639, 22279000, 22279002, 264482
1862	87372823 (3723, 3724)	Novel Protein sim. GBank gij129493jpf07313kkm_C_RABIT - MYOSIN LIGHT CHAIN KINASE, SKELETAL MUSCLE (MLCK)	Contains protein domain (PF00089) - Eukaryotic protein kinase domain	35696288, 284259, 87168474, 284369, 21908768, 284558, 284563
1863	8575037 (3725, 3726)	Novel Protein sim. GBank gij3820909jembjCAA092991 - (AJ010642) Dof protein [Drosophila melanogaster]	UNCLASSIFIED	284601, 284788, 29148627, 29148629, 284692, 284629, 284835
1864	85547832 (3727, 3728)	Novel Protein sim. GBank gij4322263jpbIAAD158631 - (AF077736) metalloproteinase CPX-1 [Mus musculus]	Contains protein domain (PF00754) - FtsH type C domain	22278999, 284259, 284807, 285018, 18108370, 284634, 284635, 284555, 284556, 284638, 18108387
1865	87740827 (3729, 3730)	Novel Protein sim. GBank gij249572759jQ33073Y258_HUMAN - HYPOTHETICAL PROTEIN KIAA0256		22278999, 284490, 29331822, 66714117, 66712502, 285006, 285007, 285008, 285009, 284591, 80433438, 285010, 285018, 284760, 284448, 284788, 29148627, 29148629, 285020, 285022, 18108385, 60432113
1868	87266816 (3731, 3732)	Novel Protein sim. GBank gij5282617jembjCABA5748_11-(AL080157) hypothetical protein [Homo sapiens]	Kinase	18108374, 284789, 18108377, 21908765, 21908768, 35696423, 56182575, 21908769, 29148629, 35696288, 35695917, 265021, 284510, 284511, 284512, 284534, 284535, 60170831, 52644150, 264555, 284891, 284259, 284558, 284692, 284557, 33657023, 60433358, 29331822, 264559, 284595, 29331827, 35696052, 33656970, 87168518, 265017, 60431602, 22279000, 264508, 284509, 18108351, 284907, 264682, 264567, 18108372, 284785, 284486
1867	84579159 (3733, 3734)	Novel Protein sim. GBank gij385930 (AF078096) - forkheadwinged helix-like transcription factor 7 [Homo sapiens]	UNCLASSIFIED	284094

1868	87357459 (3735, 3736)	Novel Protein sim. GBank g13861525femb CA93841 - (270039) cDNA EST EMBL.D32579 comes from this gene; cDNA EST EMBL.D35254 comes from this gene; cDNA EST yk224b3.5 comes from this gene; cDNA EST yk35710.5 comes from this gene [Caenorhabditis elegans]		nuclease	264489, 22278997, 22278999, 29331825, 29331628, 265006, 265009, 33657402, 8718474, 16108351, 21906765, 21906768, 21906769, 265020, 265021, 60170815, 27466264, 264626, 18108374, 264631, 18108385, 87168518, 22279000, 22279002, 264566, 264567
1869	86977292 (3737, 3738)	Novel Protein sim. GBank g14828772fep NP_004961.1 pICFA - Insulin-like growth factor binding protein, ecd labile subunit	Contains protein domain (PF00560) - Leucine Rich Repeat	glycoprotein	264506, 264509, 264908, 264906, 264909, 264910, 264591, 264600, 18108351, 264663, 264766, 264769, 35695655, 264834, 264558, 264639, 18108385, 264563, 264486
1870	95349486 (3739, 3740)	Novel Protein sim. GBank g11869859femb CAB067221 - (286089) very large tegument protein [human herpesvirus 2]		UNCLASSIFIED	28331824, 60424269, 265007, 265008, 21906754, 265017, 265018, 265019, 264288, 264766, 264686, 264686, 21806766, 21906769, 35695917, 60170815, 264692, 18108388, 35695763, 35696423, 65274791, 264638, 264639, 58528486
1871	80234464 (3741, 3742)			UNCLASSIFIED	264509, 264905, 264595, 264768, 264635, 264636, 264563, 264486
1872	80233355 (3743, 3744)	Novel Protein sim. GBank g12460316 (AF022147) - uterus-ovary specific putative transmembrane protein [Rattus norvegicus]		protease	264510, 264594, 264585
1873	80213880 (3745, 3746)				264508, 264512, 265009, 265011, 16108351, 264687, 264691, 18108370, 16108374, 264635
1874	95351136 (3747, 3748)		Contains protein domain (PF00293) - Bacterial muft protein	UNCLASSIFIED	264488, 35695817, 264239, 264905, 264907, 264908, 264909, 263976, 264511, 264635, 264636, 264637, 264836, 33657402, 264558, 18108385, 264600, 264804, 264764, 264687, 264769
1875	87330518 (3749, 3750)	Novel Protein sim. GBank g14569520 dbj AA78762.1 - (AB023155) KIAA0938 protein [Homo sapiens]		UNCLASSIFIED	35696286, 264828, 264592, 264557, 264558
1876	87112950 (3751, 3752)	Novel Protein sim. GBank g1263810 bbs122920 - collagen alpha chain [Riftia pachyptila=tube worms, Peptide. 1027 aa]		UNCLASSIFIED	264259, 29331822, 60432288, 264908, 264809, 264604, 264784, 264268, 264768, 16108376, 264556, 264558, 264559, 18108385
1877	87315206 (3753, 3754)	Novel Protein sim. GBank g13863356 b AAC83924.1 - (AF102545) riboflavin binding protein precursor [Scaphiopus couchii]		UNCLASSIFIED	264767, 264666, 264768, 263878, 264693, 264639, 265010, 264563, 264905, 264906, 264807

1878	95351058 (3755, 3756)	Novel Protein sim. GBank gll510345ibpA021434.1 - (AC006921) unknown protein (Arabidopsis thaliana)	Contains protein domain (PF01428) - Ubiquitin	Ubiquitin	264568, 264488, 35889286, 56994075, 264259, 28331822, 28331824, 28331825, 35898052, 29331828, 29146498, 264508, 264509, 264905, 284808, 264907, 264908, 264909, 264510, 264511, 264512, 285008, 264910, 33657402, 264594, 284758, 55812038, 265011, 264602, 284760, 18108351, 264762, 264681, 264862, 264764, 264389, 264288, 264786, 264687, 264786, 264789, 21908766, 35695917, 265021, 60170615, 33657023, 264692, 264693, 33657109, 27486285, 264628, 18108370, 264629, 16108374, 35696423, 264634, 264835, 264555, 264638, 264839, 63373044, 18108385, 58526406, 67168516, 264563, 264564, 264568, 264488, 264567, 264805, 264907, 264808, 265007, 264565, 264566
1879	95310883 (3757, 3758)	Novel Protein sim. GBank gll4828643bpA034082.1 (AF15184 - (AF151845) CGI-87 protein (Homo sapiens)		UNCLASSIFIED	
1880	91012878 (3759, 3760)	Novel Protein sim. GBank gll1550765jemb(CAA69283) - (Y08028) Immune associated protein 36 (Mus musculus)		UNCLASSIFIED	264766, 264691, 264692, 83373044
1881	80214949 (3761, 3762)	Novel Protein sim. GBank gll93144pjl(B40505 - Hypothetical protein - eud hepatitis 1 (strain Indiana-Funkhouser of Becker)		UNCLASSIFIED	264509, 264805, 264808, 264909, 264910, 264762, 264687, 33657023, 264632
1882	86582450 (3763, 3764)	Novel Protein sim. GBank gll2384956 (AF022985) - No definition line found (Caenorhabditis elegans)			264908, 21906768, 16106370, 283974, 87168516
1883	94216817 (3765, 3766)	Novel Protein sim. GBank gll1351218spJF47228TIES2. MOUSE - TESTIN 2 (TIES2) [CONTAINS: TESTIN 1 (TES1)]	Contains protein domain (PF00412) - LIM domain containing proteins	homeobox	264908, 264810, 67168559, 21906766, 264838

1884	95310885 (3767, 3768)	Novel Protein sim. GBank g 14928643 gb AAD34082.1 AF15184 - (AF151845) CGI-87 protein [Homo sapiens]		UNCLASSIFIED	264468, 18108394, 56161688, 22278998, 60432049, 264259, 29331622, 29331824, 29331825, 66714117, 29331627, 35696052, 29331825, 264509, 264905, 264906, 264907, 264908, 264909, 56182435, 264510, 264511, 265007, 264512, 265009, 264910, 264592, 264593, 33657402, 55612038, 264758, 55611386, 265010, 265011, 67166559, 264600, 265017, 264604, 265019, 264605, 264760, 55611150, 264761, 264682, 264763, 264683, 264764, 264286, 264369, 264766, 264688, 264768, 264769, 26148784, 35693917, 264890, 264691, 33657023, 264692, 264693, 33657109, 18108370, 264628, 60431526, 264828, 263973, 18108374, 55610764, 55811576, 35696423, 35695855, 264631, 264634, 60431850, 264637, 264636, 58162323, 264639, 18108362, 83373044, 18108365, 60432113, 22279002, 264583, 264584, 264585, 264586, 264468, 264587
1885	87644280 (3789, 3770)	Novel Protein sim. GBank g 12507155 p P37370 VPR1_YEAST - VERPROLIN		UNCLASSIFIED	56182575, 264239, 284905, 264909, 265008, 264596, 264766, 265020, 264628, 60431526, 264634, 56526486, 264080, 264563 22278996, 22278998, 60432049, 264910, 265018, 264766, 21908766, 29148629, 264890, 264693, 264628, 264555, 264486 29331822, 29331824, 29331625, 29331626, 29331827, 264906, 265007, 264661, 264766, 29148627, 264693, 18108364, 35696423, 65274791, 35885655, 264632, 56162323, 264639, 264583 263976
1886	86674082 (3771, 3772)	Novel Protein sim. GBank g 12854158 p AAC02577.1 - (AF045641) No definition line found [Caenorhabditis elegans]			
1887	94139139 (3773, 3774)	Novel Protein sim. GBank g 15174421 p NP_006023.1 PCPNE - copine VI (neuronal)	Contains protein domain (PF00168) - C2 domain	ATPase-associated	
1888	87822804 (3775, 3776)	Novel Protein sim. GBank g 13319831 p CAB10841 - (Z98046) 4L14O9.2 (Melanoma-Associated Antigen MAGE LUKE) [Homo sapiens]	Contains protein domain (PF01454) - MAGE family		
1889	91255783 (3777, 3778)	Novel Protein sim. GBank g 1083308 p p A56559 - enhancer-trap-locus-1 protein - mouse (fragment)	Contains protein domain (PF00176) - SNF2 and others N-terminal domain	helicase	
1890	67826705 (3779, 3780)	Novel Protein sim. GBank g 1240195 p BAA74876.1 - (AB020660) KIAA0653 protein [Homo sapiens]		UNCLASSIFIED	18108398, 29147620, 264807, 265009, 264800, 265018, 18108351, 264286, 264689, 21908765, 21908768, 21908769, 264691, 264692, 264693, 264628, 18108370, 264836, 264558, 264404

1891	87013695 (3761, 3762)			UNCLASSIFIED	264686, 264768, 264867, 264892, 264693, 29331622, 29331824, 264508, 264905, 294906, 18108370, 264628, 264907, 264908, 264909, 16108379, 295007, 295008, 264910, 264632, 264591, 264639, 264598, 18108364, 265010, 285011, 264801, 264805, 264563, 264389
1892	87642823 (3763, 3764)	Novel Protein sim. GBank g[1568535]cd[1568535].1 - (AB029022) KIAA1099 protein (Homo sapiens)	Contains protein domain (PF01412) - Putative GTP-ase activating protein for Arf	struct	22278895, 264508, 67168539, 18106351, 264448, 264682, 295020, 264893, 16108374, 22279000
1893	88533826 (3785, 3786)			lamnin	264569, 65274572, 22278897, 22278899, 264259, 29331822, 29331824, 66714117, 29331828, 264908, 285006, 285008, 285009, 284592, 265018, 264881, 264448, 264683, 16108354, 264359, 264684, 264865, 264788, 264887, 264889, 21906768, 265020, 265022, 60170615, 52644150, 264690, 264691, 264692, 33657023, 264893, 33657109, 264628, 18108374, 35695855, 264830, 264632, 264634, 264557, 264558, 60170394, 18108381, 18106385, 22279000
1894	86989120 (3767, 3768)				264508, 264905, 264906, 264907, 264594, 264604, 264890, 264692, 264690, 264695, 264638, 264639, 264593
1895	87631891 (3789, 3790)	Novel Protein sim. GBank g[1528257]emb[CAB5729.1] - (AL060133) hypothetical protein (Homo sapiens)	Contains protein domain (PF00435) - Spectrin repeat		56182575, 264259, 60432289, 29331828, 264107, 264805, 264808, 264910, 60170831, 264758, 265010, 265018, 264448, 264288, 264768, 33657108, 264628, 55810764, 16106379, 264834, 56182323, 56526486
1896	85673555 (3791, 3792)		Contains protein domain (PF00627) - UBA domain	UNCLASSIFIED	264807, 265008, 264862, 264686, 21906788, 264629, 264831, 264634, 264555
1897	80565569 (3793, 3794)	Novel Protein sim. GBank g[172836]SP39193ALUB_HUMAN - III ALU SUBFAMILY SP WARNING ENTRY IIII		cadherin	264259
1898	87817837 (3795, 3796)	Novel Protein sim. GBank g[1127560]SP23249MAY10_MOUSE - PROTEIN MCV-10		helicase	22278896, 22278898, 22278899, 29331824, 29331825, 60432289, 29331827, 35898052, 29331828, 265008, 265019, 264681, 264682, 264448, 264389, 52644229, 21906765, 21906766, 21906768, 21906769, 60170615, 55810764, 22279000
1899	86673097 (3797, 3798)	Novel Protein sim. GBank g[12909819] (AF031548) - erythrocyte membrane glycoprotein Rh50 (Homo sapiens)	Contains protein domain (PF00909) - Armadillo Transmembrane Family	glycoprotein	264259, 264506, 264909, 60432229, 264769, 21906765, 21906769
1900	67641858 (3799, 3800)	Novel Protein sim. GBank g[14102861] (AF017250) - vitellinogen precursor (Oreochromis aureus)		UNCLASSIFIED	264693

1901	59196647 (3801, 3802)	Novel Protein sim. GBank gll365959ispf363781S61A_RAT PROTEIN TRANSPORT PROTEIN SECC1 ALPHA SUBUNIT	Contains protein domain (PF00242) - DNA polymerase (viral) N-terminal domain	transport	264488, 52844507, 52845156, 18108396, 52848395, 52848642, 18108397, 58182575, 22278994, 22278995, 56994075, 35696286, 22278997, 22278998, 264490, 60432048, 264259, 29331822, 52845080, 29331824, 29331825, 86714117, 29331826, 60432289, 29331827, 29331828, 35696052, 35696970, 29148498, 264908, 264907, 29331830, 264908, 52844045, 264909, 264112, 265006, 284512, 265008, 264910, 265009, 60170831, 60432229, 80433356, 33657402, 80433438, 55612038, 264758, 33109954, 21908754, 33657084, 52844286, 87188474, 265010, 265011, 87188559, 265017, 265016, 265018, 18108351, 264448, 264288, 264886, 52644229, 21908765, 21906766, 21906767, 21908789, 55811957, 35695917, 265020, 285021, 52844150, 18108382, 33657023, 264693, 265987, 33657109, 33657182, 27486284, 33657349, 35695763, 18108370, 18108376, 55811576, 35698423, 35695855, 60431850, 264638, 263981, 52844332, 60170394, 83373044, 18108365, 87188518, 80432113, 264584 264107, 263978
1902	80202013 (3803, 3804)	Novel Protein sim. GBank gll44266139pAAD204511 - (AF098789) SLM-1 (Mus musculus)		dna_ma_bind	264259, 29331825, 29331827, 264508, 264807, 265008, 60170831, 60433356, 60433438, 264759, 21908754, 264448, 264288, 265021, 265022, 33657023, 264693, 55811576, 284555, 284558, 22279000
1903	87776554 (3805, 3806)	Novel Protein sim. GBank gll3747107 (AF095741) - unknown [Fetus norvegicus]		UNCLASSIFIED	264509, 264905, 264906, 264807, 264908, 265007, 264910, 264888, 264788, 264887, 264789, 264693, 264628, 18108374, 264634, 264838, 264637, 264585
1904	60434213 (3807, 3808)	Novel Protein sim. GBank gll352911ispf47147YJ80_YEAST - HYPOTHETICAL 80.2 KD PROTEIN IN CPA2-NNF1 INTERGENIC REGION		struct	264488, 264768, 264789, 284689, 29148629, 35695917, 35696286, 284259, 284692, 18108362, 33857023, 29331824, 33657109, 29146499, 264508, 264509, 264905, 264906, 264907, 66712502, 264906, 264909, 35698423, 35695855, 264510, 264511, 264512, 284910, 264834, 264635, 264637, 264638, 33657402, 264758, 85658542, 264802, 264780, 264781, 264482, 264583, 264762, 264463, 264784, 264586, 264288, 264766
1905	95351140 (3809, 3810)	Novel Protein sim. GBank gll3043714jdb BAA255211 - (AB01167) KIAA0595 protein [Homo sapiens]	Contains protein domain (PF00293) - Bacterial mult' protein		284637
1908	12763822 (3811, 3812)			UNCLASSIFIED	

1907	85351144 (3813, 3814)	Novel Protein sim. GBank gll4929505gll4929505.1 (AF15181 - (AF151816) CGI-58 protein [Homo sapiens])	Contains protein domain (PF00561) - alpha/beta hydrolase fold	hydrolase	65274572, 22278998, 35698286, 22278998, 22278998, 264258, 80432049, 29331822, 29331825, 60424268, 29331828, 35696052, 264593, 60433356, 21900754, 55811386, 85858542, 87168559, 265018, 264681, 264682, 264684, 264288, 21906765, 21906766, 21906768, 265020, 265022, 264690, 52644150, 264682, 33857023, 264693, 33657109, 35698355, 264636, 264638, 60432113
1908	95313641 (3815, 3816)	Novel Protein sim. GBank gll3986770 (AF109806) - NG22 [Mus musculus]		UNCLASSIFIED	264488, 65274572, 56182575, 56181686, 22278995, 22278998, 22278997, 22278998, 22278999, 60432049, 264258, 28331822, 60432268, 29331826, 29331827, 29331828, 29148499, 264905, 264806, 264807, 264908, 68712502, 264908, 56182435, 284510, 284511, 265008, 264910, 60170831, 264592, 60433356, 33657402, 264594, 60433438, 264596, 55812038, 33109854, 52848317, 265011, 265017, 264604, 265018, 265019, 264805, 55811150, 264681, 264448, 264288, 264886, 264688, 264769, 21906765, 21906766, 21906767, 21906768, 21906769, 55811957, 29148628, 35695917, 265020, 265022, 264691, 264692, 18108384, 65274620, 33657109, 33657348, 35695763, 18108374, 263978, 55810784, 55811576, 35698423, 65274791, 264631, 264632, 264556, 264557, 60170394, 56182323, 83373044, 18108385, 60432113, 22279000, 22279002, 264566, 264488
1909	86514505 (3817, 3818)	Novel Protein sim. GBank gll2224653dbllBA208131 - (AB002354) KIAA0356 [Homo sapiens]		UNCLASSIFIED	264259, 264508, 264905, 264906, 264907, 264908, 264511, 264910, 264593, 264758, 264764, 264766, 18108370, 264834, 264637, 264488
1910	84216821 (3819, 3820)	Novel Protein sim. GBank gll1351218sp1472281TES2.MOUSE - TESTIN 2 (TES2) [CONTAINS: TESTIN 1 (TES1)]	Contains protein domain (PF00412) - LIM domain containing proteins	homeobox	35698286, 22278998, 22278999, 35696052, 264508, 264905, 264906, 264907, 264908, 264909, 264511, 265007, 264512, 264910, 264756, 265011, 264801, 264802, 264604, 264605, 264761, 264764, 264288, 264768, 264768, 264887, 264769, 35695917, 265021, 52644150, 264692, 264628, 18108370, 264828, 18108372, 18108374, 35698555, 264631, 264634, 264635, 264636, 264637, 264638, 18108385, 264565, 264566, 264488

1811	91725345 (3821, 3822)	Novel Protein sim. GBank g14808339b PAA030184.1 AC006530 hypothetical protein [Homo sapiens]	Contains protein domain (PF01118) - DNA mismatch repair protein	nuclease	18108394, 56182575, 56182781, 29331826, 29331627, 33656970, 284906, 265007, 284591, 55812038, 87188559, 284448, 284369, 21806765, 21906768, 265022, 284891, 264893, 18108385, 55811578, 264558, 18108385, 18108388
1812	95413519 (3823, 3824)	Novel Protein sim. GBank g15688439b PAA03003.1 - (AB028974) KIAA1051 protein [Homo sapiens]	Contains protein domain (PF00098) - Zinc finger, CCHC class	UNCLASSIFIED	18108397, 56182575, 56181686, 22278994, 22278995, 58994075, 22278996, 22278999, 284259, 29331822, 29331824, 56182161, 29331825, 66714117, 35686052, 264805, 284906, 284907, 284908, 52844045, 56182435, 285007, 285008, 284910, 265008, 284591, 284596, 65274444, 55811386, 87168474, 265011, 87168559, 265018, 265019, 264760, 18108351, 264681, 264369, 264684, 264288, 264888, 264768, 21908765, 21906766, 21906767, 21906768, 21906769, 55611957, 285020, 285021, 285022, 60170615, 284692, 33657023, 284693, 18108376, 55811576, 35898423, 65274791, 284837, 58182323, 63373044, 58526466, 22279002, 284583, 284588
1813	95305546 (3825, 3826)	Novel Protein sim. GBank g1503245 NP_005865.1 pZNF2 - zinc finger protein (C2H2) homologous to mouse MOK-2			56182575, 22278994, 22278995, 56994075, 22278998, 22278999, 22278999, 29331826, 29331627, 265008, 55812038, 265010, 265017, 265018, 265019, 264681, 18108351, 264883, 284764, 264389, 264268, 284885, 264688, 264769, 21906765, 21906768, 21906769, 21906769, 55811857, 265020, 265022, 264691, 55811576, 264834, 264635, 264638, 56182323, 63373044, 18108385
1814	63423982 (3827, 3828)	Novel Protein sim. GBank g16589604 db PAA78824.1 - (AB023197) KIAA0980 protein [Homo sapiens]	Contains protein domain (PF00038) - EF hand	struct	58182575, 29331824, 35686052, 264906, 264908, 264828, 264909, 264592, 264758, 87188559, 18108351, 18108354, 264684, 264888, 33857023, 264693, 264626, 264631, 284832, 264634, 264835, 264639
1815	95340459 (3829, 3830)	Novel Protein sim. GBank g15689415 db PAA02981.1 - (AB028962) KIAA1039 protein [Homo sapiens]		UNCLASSIFIED	284259, 29331824, 29331826, 29331627, 264508, 264908, 265009, 265017, 265019, 264768, 264769, 264889, 264628, 264835, 264637, 264639, 63373044, 264585
1816	78640761 (3831, 3832)				284683, 264639

1917	87821680 (3833, 3834)	Novel Protein sim. GBank g1568539 [cdh1bAA82978.1 - (AB028950) KIAA1027 protein (Homo sapiens)]	struct	264769, 264689, 21906765, 21906768, 22278998, 264259, 264691, 264693, 29331624, 29331625, 29331626, 29331626, 264805, 264906, 264628, 264907, 264908, 264909, 264510, 264630, 264910, 264634, 264835, 264636, 264637, 264638, 263981, 264639, 264758, 18108385, 21906754, 265011, 264804, 264583, 18108351, 264782, 264763, 264566, 264784, 264768
1916	95302785 (3635, 3638)	Novel Protein sim. GBank g1528151 [gbpAAD41524.1 (AF15483) PV-1 (Rattus norvegicus)]	struct	284488, 18108382, 18108357, 21906765, 21906767, 21906768, 56182575, 21906769, 22278994, 35696288, 35695817, 22278998, 22278997, 285021, 285022, 284534, 284690, 284691, 264692, 33657023, 264893, 29331624, 29331625, 33657109, 29331626, 52645129, 35696052, 29331628, 27486262, 27486284, 35695763, 264508, 284905, 264509, 264908, 264628, 264907, 18108370, 264808, 264629, 264809, 18108372, 18108374, 263978, 35696423, 35695855, 264510, 264511, 265008, 265007, 264512, 265006, 264831, 265009, 264810, 264634, 264635, 264555, 264636, 264556, 264637, 264557, 264593, 264638, 264594, 60170394, 264595, 264596, 264598, 83373044, 264758, 52846317, 16108385, 52644298, 58526468, 87188518, 285010, 265011, 87188559, 264600, 264601, 264602, 265017, 264603, 264604, 265018, 264605, 264780, 264781, 264482, 264564, 18108351, 264782, 284662, 264585, 264448, 264764, 264566, 264488, 264567, 264369, 264288, 264786, 264487, 264685
1919	84143847 (3637, 3638)	Novel Protein sim. GBank g13878564 [embhCAB012371 - (Z77667) cDNA EST EMBL:CO8125 comes from this gene; cDNA EST EMBL:CO9753 comes from this gene (Caenorhabditis elegans)]	oxidase	22278997, 29331622, 265007, 60170831, 60432229, 80433438, 264448, 264682, 264286, 55811957, 33657023, 33657109, 85274791, 56182323, 22278902
1920	91228953 (3839, 3840)	Novel Protein sim. GBank g11809231 (AC000115) - coded for by human cDNAs R76043 (NID:9850725), R56857 (NID:963495) and H12888 (NID:9877689) (Homo sapiens)	UNCLASSIFIED	264510, 264511, 264566
1921	79555228 (3841, 3842)	Novel Protein sim. GBank g14580997 [gbpAAD24571.1 (AF12108) cAMP Inducible 2 protein (Mus musculus)]	UNCLASSIFIED	264693

1922	87641863 (3843, 3844)	Novel Protein sim. GBank gll136595ippp20451vtr2 - CHICK - VITELLOGENIN II PRECURSOR (MAJOR VITELLOGENIN) [CONTAINS: LIPOVITELLIN I (LVJ), PHOSVITIN (PV), LIPOVITELLIN II (LVII), YGP40]		UNCLASSIFIED	264686, 264688, 264490, 16108370, 264909, 16106374, 265008, 264557, 264594, 16108351
1923	84323589 (3845, 3846)	Novel Protein sim. GBank gll119110ispf03211EBN1_EBV -EBNA-1 NUCLEAR PROTEIN		UNCLASSIFIED	264488, 264489, 22278995, 264094, 264258, 35698052, 264508, 264905, 264908, 264907, 264908, 66712502, 264908, 264511, 264512, 265007, 264910, 265008, 264591, 264592, 264593, 264594, 264595, 264758, 264596, 264759, 265011, 265017, 265018, 265019, 55811150, 264681, 264782, 264448, 264784, 264288, 264389, 264786, 264787, 264886, 264687, 264788, 264789, 56161562, 264689, 21908766, 264691, 33857023, 264693, 65274620, 33857109, 18108370, 264628, 264629, 35698423, 264630, 264631, 264632, 264634, 264555, 264635, 264637, 264638, 264639, 264556, 63373044, 67168518, 264583, 264584, 264585, 264586, 264587 29331826, 264906, 264908, 264595, 265017, 265018, 265019, 265021, 264691, 264693, 264637, 18108385, 264565
1824	67338925 (3847, 3848)	Novel Protein sim. GBank gll3877655fembCAA86571 - (Z72511) possible zinc finger protein. cDNA EST EMBL:M8915 comes from this gene; cDNA EST EMBL:D71533 comes from this gene; cDNA EST EMBL:D72314 comes from this gene; cDNA EST EMBL:D75164 comes from this gene; cDNA EST EMBL:C1...	Contains protein domain (PF00097) - Zinc finger, C3HC4 type (RING finger)		
1925	67628338 (3849, 3850)	Novel Protein sim. GBank gll4981903ipbAAC36415, 1AED0176 - (AEC01789) ribosomal protein S15 (Thermotoga maritima)	Contains protein domain (PF00312) - Ribosomal protein S15	ribosomalprot	22278995, 22278996, 22278997, 264259, 29331824, 68714117, 29148499, 264909, 52644045, 265008, 265009, 264758, 265011, 265017, 264605, 264448, 264288, 264892, 33857109, 18108374, 80170394 264905, 264906, 264907, 264910, 264593, 265018, 264780, 264784, 264286, 264892, 294893, 263978, 264631, 264634, 264837, 264563 264689, 264631
1926	88094739 (3851, 3852)	Novel Protein sim. GBank gll2246532 (U93672) - ORF 73, contains large complex repeat CR 73 [Kaposi's sarcoma- associated herpesvirus]		stud	
1927	85654857 (3853, 3854)	Novel Protein sim. GBank gll5043632ipbBA254801 - (A8011126) KIAA0554 protein [Homo sapiens]		int	264489, 264259, 265017, 265021, 264692
1928	87788054 (3855, 3856)	Novel Protein sim. GBank gll1665761ipbBA1133771 - (DB7433) KIAA0246 [Homo sapiens]	Contains protein domain (PF00193) - Extracellular link domain		264506, 264591, 33857402, 265017, 264768, 264632, 264596, 264639
1929	88897236 (3857, 3858)	Novel Protein sim. GBank gll5001983ipbAACD37247, 1AF13432 - (AF134321) chimeric AF-GP/tryptophan-like serine protease precursor [Discotheca mawsonii]		UNCLASSIFIED	

1930	67688128 (3859, 3860)	Novel Protein sim. GBank 6l1708230spIP52663INBL4_MOUSE - NBL4 PROTEIN		phosphatase	35696286, 29331626, 264605, 264907, 264906, 264909, 264511, 264910, 264756, 264601, 265017, 265019, 264605, 264760, 264764, 264766, 264668, 264768, 265022, 35698423, 264638, 60432113
1931	67767279 (3861, 3862)	Novel Protein sim. GBank gll4104634 (U01840) - serine/threonine kinase [Mus musculus]	Contains protein domain (PF00066) - Eukaryotic protein kinase domain	kinase	264306, 264906, 60432229, 264756, 264764, 264288, 265020, 264692, 264634, 264637 264684, 264691, 264635
1932	15030872 (3863, 3864)			UNCLASSIFIED	264595
1933	11613668 (3865, 3866)			struct	56162575, 56162435, 264510, 264757, 264758, 55612036, 55611366, 265016, 55611150, 21806765, 264691, 264631, 264635, 264637
1934	84426380 (3867, 3868)	Novel Protein sim. GBank gll4115748jdbPAA38484 - (AB022023) nonmuscle myosin heavy chain B [Bos taurus]		UNCLASSIFIED	264688, 265011, 264511, 264905, 18108351, 264564, 264681, 264259, 16108370, 264568, 264764, 264369, 264595
1935	67752511 (3869, 3870)			UNCLASSIFIED	60432286, 265007, 265010, 265011, 265019, 33657106, 16108374
1936	65414338 (3871, 3872)	Novel Protein sim. GBank gll4827040jeflNP_005110_1tpTRAP - thyroid hormone receptor-associated protein, 150 kDa subunit	Contains protein domain (PF00023) - kinase		85656542, 21806767, 35685817, 60170615, 264693, 33657109
1937	84647141 (3873, 3874)	Novel Protein sim. GBank gll543167jplfIS3771 - ankyrin, erythrocyte - mouse	Ank repeat Contains protein domain (PF00889) - collagen		264468, 26146488, 264805, 264559
1938	67403277 (3875, 3876)	Novel Protein sim. GBank gll4544431gblpAAD2340_1AC00695 - (AC006955) hypothetical protein [Arabidopsis thaliana]	Culin family		
1939	81004878 (3877, 3878)	Novel Protein sim. GBank gll500858jdbPAA032101 - (D14166) 50kDa lectin [Bombyx mori]		UNCLASSIFIED	65274572, 22278889, 60432288, 56182435, 60433356, 265017, 21806765, 21806766, 21806768, 55611957, 27466284, 35696423, 60432113, 264564
1940	67346810 (3879, 3880)	Novel Protein sim. GBank gll1946300jembCAA731321 - (Y12529) hypothetical protein [Silene latifolia]	Contains protein domain (PF00560) - Leucine Rich Repeat	struct	264466, 26331622, 264446, 264663, 264286, 265020, 33657023, 264631
1941	84141717 (3881, 3882)	Novel Protein sim. GBank gll4208386 (AF060570) - ng-1 protein [Mus musculus]		UNCLASSIFIED	56694075, 22278898, 264259, 26331624, 29331827, 264905, 265008, 33657084, 265017, 265016, 264288, 264687, 21806765, 21806766, 21806767, 265020, 52644150, 27466284, 83373044, 16108387, 60432113, 22279002, 264565
1942	87641670 (3883, 3884)	Novel Protein sim. GBank gll4927204jgblpAAD33049_1AF13391 (1) ARL-6 interacting protein-4 [Mus musculus]		UNCLASSIFIED	264466, 16108396, 29331825, 27466281, 264506, 16108370, 16108374, 264462
1943	84325288 (3885, 3886)	Novel Protein sim. GBank 6l13122852sp10157391TPD_DICD1 - TPD PROTEIN	Contains protein domain (PF00400) - WD domain, G-beta repeat	kinase	22278668, 29331822, 29331827, 35686052, 264511, 265006, 264592, 60432226, 265017, 265018, 265019, 264684, 264692, 33657108, 65274791, 264636

1944	84232858 (3887, 3888)	Novel Protein sim. GBank g11799570(dblBA13432) - (D87671) TTP120 [Rattus norvegicus]		UNCLASSIFIED	65274572, 22278994, 22278985, 22278996, 22278999, 22278999, 264259, 52645080, 29331824, 29331828, 29331827, 66712502, 56182435, 264512, 264910, 60170831, 60432229, 60433358, 33857402, 60433438, 284598, 33109954, 21906754, 67168474, 87168558, 265017, 265018, 265019, 18108351, 284368, 284888, 264768, 21906765, 21906768, 21906767, 21906789, 35695917, 265020, 265021, 60170615, 284692, 33657023, 16108370, 18108374, 35695423, 35695655, 264634, 60170394, 264639, 83373044, 18108385, 56526488, 67168518, 60432113, 22279000, 284563, 26468, 22278998, 264510, 284511, 18108351, 264683, 284468, 264587
1945	87641872 (3888, 3890)	Novel Protein sim. GBank g14927204(gb)AAD3049.1(A-F13391 - (A-F13391-1) ARL-6 interacting protein-4 [Mus musculus])		UNCLASSIFIED	26468, 22278998, 264510, 284511, 18108351, 264683, 284468, 264587
1948	87443890 (3891, 3892)	Novel Protein sim. GBank g12498104(sp)Q27869(AAD30, BOVIN - ADRENAL MEDULLA 50 KD PROTEIN		UNCLASSIFIED	60432289, 29331827, 35696052, 265007, 285008, 60433356, 60433436, 264389, 56181562, 21906787, 52644150, 264693, 27488264, 284637, 87168518, 264563
1947	86438682 (3893, 3894)	Novel Protein sim. GBank g13914801(sp)Q54688(RP2, RAT - DNA-DIRECTED RNA POLYMERASE 1 135 KD POLYPEPTIDE (RNA POLYMERASE 1 SUBUNIT 2) (RPA135) (RNA POLYMERASE 1 127 KD SUBUNIT)		mapolymerase	22278998, 264905, 264906, 264908, 264909, 264512, 264758, 264762, 264682, 284883, 264764, 264286, 264788, 284686, 21906768, 264693, 18108374, 35695855, 264635, 264637, 284638, 18108385, 22279002
1946	85189174 (3895, 3896)	Novel Protein sim. GBank g15420387(emh)CAB46678.11 - (AJ243459) proteophosphoglycan [Leishmania major]		struct	264909, 60170831, 264591, 264594, 255010, 265011, 264764, 264369, 284889, 284631, 264638
1949	7640129 (3897, 3898)				264389
1950	87788531 (3899, 3900)	Novel Protein sim. GBank g1387676(emh)CAA93465.11 - (Z69637) predicted using GeneRinder. Similarity to E. coli hypothetical protein YCAC (SW:YCAC_ECOLI) [Caenorhabditis elegans]		UNCLASSIFIED	264488, 264768, 264688, 264688, 264258, 29331822, 33657109, 29331828, 264103, 264509, 16108370, 35695855, 264510, 265008, 265009, 33657402, 18108385, 265018, 264563, 55811150, 18108351, 264369, 264288, 16108354
1951	86868253 (3901, 3902)	Novel Protein sim. GBank g12626753(dbl)BA234241 - (AB008782) sulfate transporter [Arabidopsis thaliana]		transport	56182575, 22278997, 52645080, 29331824, 29331825, 29331827, 55812038, 52648317, 265018, 265019, 284369, 21906765, 21906767, 55611857, 265020, 265021, 33657023, 264683, 35695763, 56182323, 22279002
1952	87088775 (3903, 3904)	Novel Protein sim. GBank g14928833(gb)AAC34077.1(A-F15184 - (A-F15184) CG1-62 protein [Homo sapiens]		reductase	264581, 264593, 264594, 264595, 264555, 264558, 264557, 284558, 284585

1953	20470371 (3905, 3906)	Novel Protein sim. GBank gi1186715sp P31721C1QB_RAT - COMPLEMENT C1Q SUBCOMPONENT, B CHAIN PRECURSOR	Contains protein domain (PF00443) - C1q domain	complement	264259, 264558
1954	91228025 (3807, 3908)	Novel Protein sim. GBank gi1240271db BA74914.1 - (A6020898) KIAA0891 protein [Homo sapiens]	Contains protein domain (PF00443) - Ubiquitin carboxy-terminal hydrolase family 2	ubiquitin	264468, 263394, 18108394, 35696288, 22278998, 29331822, 68714117, 29331826, 29331827, 264508, 264509, 264905, 264906, 264907, 264908, 264909, 264510, 265008, 265007, 265008, 264910, 285008, 80170831, 60432229, 265011, 265017, 264603, 264604, 265019, 18108351, 284685, 264768, 264769, 35695917, 265020, 265021, 264691, 33657023, 284692, 33657109, 264628, 18108374, 35696423, 55811576, 35695855, 264630, 264635, 264636, 264555, 264558, 264638, 264557, 264639, 264558, 264559, 83373044, 18108385, 264563, 264564, 264566, 264466, 264567
1855	93308310 (3909, 3910)			UNCLASSIFIED	56182575, 56181686, 22278994, 22278999, 264259, 29331822, 58182181, 28331824, 29331825, 29331826, 29331827, 35696052, 264508, 29331830, 265008, 285009, 264591, 55812038, 87168474, 265017, 265018, 265019, 264448, 264766, 21906765, 21906768, 21906767, 55811957, 265020, 285021, 52645129, 33657109, 27486284, 33657349, 35695763, 60431528, 18108374, 55811576, 35695855, 284635, 60431850, 264639, 83373044
1956	93092121 (3911, 3912)	Novel Protein sim. GBank gi1865821db BAA134071 - (D87489) Similar to D-melanogaster cadherin-related tumor suppressor [Homo sapiens]	Contains protein domain (PF00028) - cadherin	cadherin	264488, 58182575, 22278996, 22278999, 264259, 29331822, 28331824, 29331825, 29331827, 35696052, 29331828, 264508, 264908, 264908, 264909, 264511, 285007, 264910, 284591, 55812038, 265010, 285018, 18108351, 264768, 56181582, 264689, 21908788, 21908789, 265022, 264691, 264628, 18108374, 55810764, 55811576, 35695855, 264631, 264632, 264635, 264637, 264639, 80170394, 58182323, 83373044, 18108385, 22279000, 22279002, 264583, 264584

1957	94326510 (3913, 3914)	Novel Protein sim. GBank gik458987[idb]BAA76856.11 - (AB023229) KIAA1012 protein [Homo sapiens]		UNCLASSIFIED	52648842, 56182575, 22278997, 22278998, 22278999, 29331824, 66714117, 29331827, 29146498, 264593, 33857402, 33109954, 87188474, 265018, 264448, 264369, 264286, 264786, 21906766, 21906767, 21906768, 21906769, 265020, 265021, 264692, 65274620, 27466264, 33657349, 27486265, 35695855, 22279002, 284462
1956	95313902 (3915, 3916)	Novel Protein sim. GBank gik4240227[idb]BAA74892.11 - (AB020676) KIAA0869 protein [Homo sapiens]		UNCLASSIFIED	22278998, 264092, 264094, 264259, 60432049, 29331824, 56182181, 66714117, 264107, 264109, 264909, 264511, 60170831, 60432229, 21906754, 265010, 21906769, 35695917, 265022, 65274620, 263987, 263976, 35696423, 284631, 264632, 264634, 264635, 18106385, 22279000, 22279002, 264593, 265019
1959	65701470 (3917, 3918)	Novel Protein sim. GBank gik2261983[idb]BAA10860.1 - (Z98056) hypothetical protein [Schizosaccharomyces pombe]		ubiquitin	264805, 264806, 264907, 264908, 264909, 265006, 265007, 264910, 264595, 265017, 264604, 265018, 18106351, 264784, 264369, 264766, 264768, 21906765, 18106366, 264629, 18106378, 264635, 264636, 264637, 264638, 264486
1960	80308608 (3919, 3920)	Novel Protein sim. GBank gik2274651[idb]BAA21515.1 - (DB4159) 3-7 gene product [Homo sapiens]		struct	264635
1981	16292807 (3921, 3922)			UNCLASSIFIED	65274572, 264592, 264593, 265019, 264691, 65274572, 18106398, 35696266, 29331825, 60432269, 29331827, 264826, 265006, 265008, 60433356, 60433438, 21906754, 265020, 265021, 33657023, 33657109, 27486285, 35695855, 264555
1982	91006385 (3923, 3924)	Novel Protein sim. GBank gik3721653[idb]BAA33581.1 - (AB012933) acyl-CoA synthetase 5 [Rattus norvegicus]		epn	264488, 264092, 264259, 264509, 264905, 264906, 264907, 264908, 264909, 264510, 264511, 265007, 265009, 264910, 264502, 264593, 264594, 264595, 264756, 264600, 264603, 264604, 264605, 264760, 264762, 264448, 264764, 264288, 264685, 264766, 264768, 264769, 21906766, 264691, 264692, 264693, 16108370, 264628, 264629, 18106374, 264630, 264631, 264634, 264636, 264637, 264638, 18106382, 6337044, 18106385, 264463, 264584, 264585, 264586, 264488, 264567
1983	80836017 (3925, 3926)	Novel Protein sim. GBank gik5626263[idb]CABA5757.11 - (AL080169) hypothetical protein [Homo sapiens]		cathecin	264488, 264510, 264760, 264768, 264486
1984	84317605 (3927, 3928)	Novel Protein sim. GBank gik4107077[idb]BAA36294.1 - (AB001773) PEM-8 [Clona savigny]		ubiquitin	
1985	94317445 (3929, 3930)	Novel Protein sim. GBank gik4107077[idb]BAA36294.1 - (AB001773) PEM-8 [Clona savigny]	Contains protein domain (PF01426) - AN1-like Zinc finger		

1988	94192058 (3931, 3932)	Novel Protein sim. GBank gll4929707jbb/AADD4114; [AF15187 - (AF15187) CGI-119 protein [Homo sapiens]	Contains protein domain (PF01027) - Uncategorized protein family	[glycoprotein	22276999, 264092, 264259, 29331826, 29331826, 29146498, 264595, 265011, 264448, 18108354, 264286, 264894, 264788, 264665, 264668, 265022, 264891, 264692, 18108370, 18108377, 264555, 18108381, 18106385, 264488, 264567
1987	87396123 (3933, 3934)	Novel Protein sim. GBank gll2957270 (AF044576) - phospholipase C PLC210 [Caenorhabditis elegans]	Contains protein domain (PF00388) - Phosphatidylinositol-specific phospholipase C, X domain	esterase	29331824, 265010, 265017, 264288, 21905784, 263981, 56528488
1986	68095641 (3935, 3936)	Novel Protein sim. GBank gll2564953 (AF030601) - unknown [Mus musculus]	Contains protein domain (PF00008) - EGF-like domain	oncogene	35696286, 264805, 264508, 264806, 264907, 264908, 264909, 264511, 264512, 285006, 264810, 265009, 264594, 264757, 264756, 264604, 264605, 264760, 264762, 264682, 264784, 264685, 264768, 264787, 264889, 264691, 264683, 264628, 264628, 35698423, 35695855, 264631, 264632, 264634, 264635, 264636, 264637, 18108380, 264564, 264565, 264566, 264567
1989	84328529 (3937, 3938)	Novel Protein sim. GBank gll2911274 (U20328) - epidion 1		UNCLASSIFIED	22278985, 22278996, 35696052, 264508, 264908, 18108351, 264482
1970	80598049 (3939, 3940)	[Nephila clavipes] Novel Protein sim. GBank gll4050087 (AF109907) - S164		UNCLASSIFIED	264908, 264288, 264766, 264636
1971	94843914 (3941, 3942)	[Homo sapiens] Novel Protein sim. GBank gll134208sp/p059593SANT_PLAFV - S-ANTIGEN PROTEIN PRECURSOR		collagen	264468, 264489, 22278996, 264259, 60432049, 68741117, 28331828, 60432289, 29331827, 35696052, 264508, 264805, 284509, 284908, 284807, 284908, 264909, 264510, 284511, 264512, 264910, 264591, 264592, 60432229, 60433358, 264595, 264588, 264800, 264804, 264605, 264760, 18108351, 264448, 264764, 264288, 264766, 264768, 264769, 21908765, 33657023, 264692, 18108370, 284629, 35696423, 69274781, 35695855, 264632, 264635, 264555, 284838, 264637, 264638, 264639, 18108385, 60432113, 22278000, 264583, 264584, 264585, 264586, 264486
1972	87645444 (3943, 3944)	Novel Protein sim. GBank gll519623jdbj/BA75871.11 - (AB017618) homologous to the yeast YGR163 gene [Mus musculus]	Contains protein domain (PF01462) - Leucine rich repeat N-terminal domain		22278899, 284259, 28331822, 56182181, 60432289, 29331827, 52644045, 284909, 285006, 264511, 285008, 52644298, 285016, 285018, 264781, 264888, 21908768, 21908769, 264691, 264893, 33857109, 33857182, 264556, 52644332, 264558, 60432113
1973	88395533 (3945, 3946)			UNCLASSIFIED	28331826, 264692, 35698423, 264831, 264555, 264556, 264557, 264558, 264559
1974	80396829 (3947, 3948)	Novel Protein sim. GBank gll3309543 (AF036382) - MLL [Fugu rubripes]		UNCLASSIFIED	264682, 264764, 264563

1875	94316479 (3949, 3950)			UNCLASSIFIED	264488, 66714117, 29331826, 29331828, 56182435, 265008, 264757, 55812036, 265010, 265017, 264369, 55811957, 65274791, 35895855, 58182323, 80432113, 264259, 35696052, 265018, 265020, 265021, 33657109, 56526466, 264908, 264596, 265021, 264566
1878	95356914 (3951, 3952)			UNCLASSIFIED	
1877	84652664 (3953, 3954)	Novel Protein sim. GBank g 249526 sp C07782 MASU_PAT - SODIUM/SULFATE COTRANSPORTER (NA(+)/SULFATE COTRANSPORTER)		homeobox	
1878	87447845 (3955, 3956)	Novel Protein sim. GBank g 103421 p A33471 - transcription factor NTF1 - fruit fly (Drosophila melanogaster) (fragment)		transcription factor	60170631, 264568
1879	87827708 (3957, 3958)	Novel Protein sim. GBank g 2244615 emb CAB10238.11 - (297336) hypothetical protein [Arabidopsis thaliana]		UNCLASSIFIED	29331826, 29146496, 264905, 264907, 265007, 265009, 265010, 265018, 264868, 16108359, 21906768, 35695917, 265020, 60170615, 264693, 16108368, 16108370, 264631, 264635, 264558, 264558, 16108364, 22279000, 264585
1880	68577058 (3959, 3960)	Novel Protein sim. GBank g 4758280 ref NP_004842.1 p USP1 - Ubiquitin carboxyl-terminal hydrolase, X-linked		ubiquitin	264468
1881	87606974 (3961, 3962)	Novel Protein sim. GBank g 14406893 p A020060 - (AF131848) Unknown [Homo sapiens]		UNCLASSIFIED	22276995, 22276997, 22276998, 264259, 29331825, 29331828, 29146486, 29146499, 264107, 264908, 264910, 264595, 21906754, 265010, 265017, 265018, 265019, 264448, 264288, 21906767, 33657023, 27486284, 18108370, 16108374, 16108376, 264830, 264631, 264635, 16108385, 67168518, 22279000, 264482, 264564
1882	80885367 (3963, 3964)	Novel Protein sim. GBank g 15689523 p BAA33045.1 - (AB029016) KIAA1093 protein [Homo sapiens]			65274572, 29331825, 35686052, 33656970, 264909, 265008, 55811386, 264780, 264666, 264691, 27486284
1883	95096866 (3965, 3966)	Novel Protein sim. GBank g 13417297 (AC002310) - Unknown gene product [Homo sapiens]	Contains protein domain (PF00096) - Zinc finger, C2H2 type	transcription factor	22276998, 35696286, 264256, 29331822, 20281099, 29331824, 60432289, 29331827, 264509, 264905, 264906, 264907, 68712502, 264808, 52644045, 264909, 264510, 264512, 264910, 265009, 264591, 264582, 60433356, 60433438, 264756, 265010, 264800, 264803, 264804, 264780, 264782, 264783, 264784, 264786, 264687, 264788, 264789, 21908765, 55811957, 35695917, 264890, 264892, 264693, 264628, 264829, 263978, 18108378, 35696423, 35695855, 20281071, 264632, 264634, 264635, 264555, 264638, 264637, 264638, 264639, 63373044, 16108385, 22279000, 22279002

1984	85760969 (3967, 3968)	Novel Protein sim. GBank gij2896895femb[CAAT1714, 11 - (AL021687) fadD14 [Mycobacterium tuberculosis]		synthase	264686, 21906766, 55611957, 56894075, 265020, 265021, 22278999, 265022, 264259, 29331622, 33657182, 29146499, 264626, 18108370, 264908, 264629, 55811576, 35695655, 265008, 265007, 264591, 21906754, 33857084, 265010, 265017, 265019, 264286
1985	85636897 (3968, 3970)	Novel Protein sim. GBank gij5712131gplACD47379.1(AF12049 - (AF120499) DEM1 protein [Homo sapiens])		glycoprotein	264760, 264286, 263378, 55811576, 264637, 56162323, 18106365, 264564
1986	60200507 (3971, 3972)	Novel Protein sim. GBank gij5712131gplACD47379.1(AF14457 - (AF144573) Mx-interacting protein Kinase PKM [Mesocricetus auratus]	Contains protein domain (PF00069) - Eukaryotic protein kinase domain	UNCLASSIFIED	264486, 264629
1987	87011117 (3973, 3974)	Novel Protein sim. GBank gij4666443gplAACD31310.1(AF14457 - (AF144573) Mx-interacting protein Kinase PKM [Mesocricetus auratus]		UNCLASSIFIED	22276999, 29331830, 265007, 265018, 21906768, 33657023, 264692, 264693, 18108377, 264635, 60170394, 22278002
1988	94122106 (3975, 3976)	Novel Protein sim. GBank gij2601701 (AF042379) - spindle pole body protein spc97 homolog CCR2 [Homo sapiens]		ubulin	264905, 264906, 264907, 264908, 264909, 264910, 264591, 264593, 264756, 264764, 264686, 264768, 265021, 264692, 264626, 264629, 35695655, 264630, 264635, 264636, 264637, 264638, 264639, 264483
1989	91225225 (3977, 3978)	Novel Protein sim. GBank gij2601701 (AF042379) - spindle pole body protein spc97 homolog CCR2 [Homo sapiens]			60432049, 60432288, 52644045, 56162435, 264112, 265007, 33657402, 52644229, 21906765, 21906766, 21906769, 55811957, 33657023, 263967, 33657109, 18108370, 22279000, 22279002
1990	65689686 (3979, 3980)	Novel Protein sim. GBank gij5701727[idb]BAAY4822.1 - (AB024728) alpha-1,3-D-mannoside beta-1,4-N-acetylglucosaminyltransferase IV-homolog [Homo sapiens]			264506, 264757, 264764, 18108381
1991	85353114 (3981, 3982)	Novel Protein sim. GBank gij4240267[idb]BAAY4822.1 - (AB020706) KIAA0899 protein [Homo sapiens]	Contains protein domain (PF01602) - Adipon N terminal region	glycoprotein	18108394, 56182575, 22278894, 35696286, 56894075, 22278997, 22278899, 29331622, 29331824, 29331825, 60432289, 29331628, 264506, 264906, 264907, 264908, 56162435, 264510, 265007, 21906754, 33109954, 67166474, 265017, 265018, 265019, 264762, 18108351, 264763, 264683, 264368, 264286, 264685, 264766, 264687, 264769, 21906765, 21906768, 21906769, 55811957, 265020, 60431526, 263974, 18108379, 35695655, 264555, 264557, 264639, 83373044, 18108384, 67166518, 60432113, 22279000, 22279002, 264564, 264488
1992	95317232 (3983, 3984)	Novel Protein sim. GBank gij2246532 (U93872) - CRF 73, contains large complex repeat CR 73 [Kaposi's sarcoma-associated herpesvirus]	Contains protein domain (PF00096) - Zinc finger, C2H2 type	struct	29331827, 264906, 264907, 264908, 265007, 264603, 264766, 264686, 264768, 21906788, 264628, 264635, 264636, 18108365, 56526466, 264566, 264567
1993	80054763 (3985, 3986)	Novel Protein sim. GBank gij2665091 (U80761) - CTG26 alternate open reading frame [Homo sapiens]		UNCLASSIFIED	264592, 35696423

1994	84329114 (3987, 3988)	Novel Protein sim. GBank g15630077igpAAD5822;1pAC00601 - (AC006017) similar to ALR; similar to AAC51735 (P1Dg2358287) [Homo sapiens]	Contains protein domain (PF00856) - [nucleopolymerase SET domain]	UNCLASSIFIED	264488, 22278997, 22278998, 22278999, 29331822, 29331824, 29331825, 29331826, 29331827, 29331828, 284905, 284907, 284908, 265008, 604332229, 33657402, 60433356, 264757, 60433438, 284758, 33103954, 285011, 265017, 285018, 265019, 264684, 264369, 264685, 264688, 264788, 21908765, 21908787, 21908788, 21908789, 265020, 264690, 18108370, 284635, 284555, 55274620, 18108370, 284635, 284555, 264556, 264557, 56182323, 83373044, 56526486, 22279000, 22279002, 264564 284488, 18108388, 22278994, 56984075, 22278996, 33688286, 22278997, 22278999, 284258, 29147620, 58182181, 29331824, 604332289, 29331828, 29331827, 33696052, 29148499, 284905, 264907, 66712502, 58182435, 285008, 285007, 285008, 285009, 60431735, 60433356, 33657402, 284585, 55812038, 33857084, 55811388, 85858542, 265010, 265011, 265017, 265018, 265019, 264761, 264762, 284448, 284683, 284764, 264288, 264766, 284688, 284788, 264769, 56181562, 284689, 21908765, 21908788, 21908789, 28148627, 21908768, 21908769, 29148629, 29148784, 265020, 285021, 284690, 18108361, 264683, 27486262, 27486264, 27488285, 18108370, 60431528, 18108374, 18108377, 33696423, 55811578, 65274791, 33696585, 284631, 264834, 264835, 284555, 284836, 60431850, 284557, 264558, 284559, 83373044, 2078451, 87188518, 284404, 80432113, 284567 264564
1995	85414353 (3989, 3990)	Novel Protein sim. GBank g14827040refNP_005110;1p1TRAP - thyroid hormone receptor-associated protein, 150 kDa subunit		UNCLASSIFIED	
1998	80234188 (3991, 3992)	Novel Protein sim. GBank g1791148fembCAA600201 - (X86028) extensin-like protein (Myra unguiculata)		UNCLASSIFIED	
1997	87028423 (3993, 3994)	Novel Protein sim. GBank g12642034 (AF034547) - protein phosphatase M130 myosin binding subunit (Ovis aries)	Contains protein domain (PF00023) - phosphatase Ank repeat		264908, 264909, 264592, 264593
1998	85282704 (3995, 3996)	Novel Protein sim. GBank g14589834clb1BA476839.11 - (AB023212) KIAA0995 protein (Homo sapiens)		UNCLASSIFIED	284113, 284085, 284555, 284567

1989	94324903 (3997, 3998)	Novel Protein sim. GBank g15225312igbLAD-00646; 1AF07244 - (AF072441) calcineurin binding protein cabin 1 (Homo sapiens)	Contains protein domain (PF00515) - UNCLASSIFIED TPR Domain	UNCLASSIFIED	18108394, 18108397, 35896286, 60424289, 29331827, 29331828, 35896052, 285008, 264512, 56811386, 265010, 265018, 265019, 55811150, 18108351, 284783, 284882, 264369, 284665, 284688, 58181582, 265020, 264691, 33657023, 264693, 33657109, 27486264, 18108370, 18108379, 35895855, 264634, 284635, 284636, 284555, 284557, 56182323, 18108362, 284559, 83373044, 60432113, 22279000, 284563, 284564, 264568
2000	95413705 (3999, 4000)	Novel Protein sim. GBank g11723232isp10155YATA_SCHPO - HYPOTHETICAL 90.8 KD PROTEIN CID4.10 IN CHROMOSOME 1		UNCLASSIFIED	52646365, 52646842, 22278994, 22278995, 22278998, 35896286, 22278997, 22278998, 22278999, 264259, 52845080, 29331822, 28331828, 29331827, 29331828, 35896052, 264108, 29331830, 5284045, 285007, 285008, 60170631, 284592, 284593, 33657402, 60433438, 21908754, 52644286, 265017, 265018, 265019, 264761, 284389, 264286, 52844229, 21908785, 21908786, 21908787, 21908788, 21908789, 35895917, 265020, 265021, 60170815, 52644150, 33657023, 65274620, 52845129, 27486261, 27486282, 27486284, 35895763, 35896423, 35895855, 264631, 52644332, 56182323, 60170394, 83373044, 56528468, 22279002, 264566, 284567
2001	95072534 (4001, 4002)	Novel Protein sim. GBank g107560p1B38637 - Ras inhibitor (clone JC265) - human (fragment)		UNCLASSIFIED	264789, 52844229, 65274572, 21806768, 22278996, 35896286, 35895817, 265020, 22278999, 284534, 284490, 284258, 284692, 60432288, 33657109, 35896052, 284508, 264508, 18108370, 60431528, 18108374, 35896423, 65274791, 35895855, 284510, 264511, 284512, 265009, 284834, 284636, 284556, 264556, 284836, 284557, 284558, 264558, 60433438, 83373044, 284759, 18108385, 285011, 284800, 264801, 60432113, 284603, 284604, 264605, 264448, 264288, 284785
2002	80236368 (4003, 4004)	Novel Protein sim. GBank g1729433ispP38657IER60_BOVIN - PROBABLE PROTEIN DISULFIDE ISOMERASE ER-60 PRECURSOR (ERP60) (58 KD MICROSOMAL PROTEIN) (P58)	Contains protein domain (PF00085) - isomerase Thioredoxin		264907, 285008, 264810, 264603, 284892, 264828, 18108374, 264558, 264557
2003	80074448 (4005, 4006)	Novel Protein sim. GBank g165389p1A27040 - neurofilament triplet M protein - chicken (fragment)		UNCLASSIFIED	284805, 284806, 284806, 284810, 264586, 285017, 18108351, 284892, 284828, 284634, 264566

2004	95317318 (4007, 4008)	Novel Protein sim. GBank g1488424[emb]CA843230.11 - (AL049998) hypothetical protein [Homo sapiens]	Contains protein domain (PF00076) - RNA recognition motif (a.k.a. RRM, RBD, or RNP domain)	dna, me, bind	52645156, 52646842, 52646365, 58182575, 22278994, 22278995, 5694075, 22278998, 35698286, 22278997, 22278998, 22278999, 60432049, 264259, 52845080, 29331824, 29331828, 29331827, 35698052, 29331828, 35696970, 29331830, 264908, 264592, 60433358, 33657402, 52846317, 21906754, 33657084, 52644298, 87186474, 87186559, 265017, 285016, 265019, 264783, 264683, 264286, 52644228, 21906765, 21906768, 21906787, 21906768, 21906769, 35695917, 265020, 265021, 265022, 52644150, 33657023, 52845129, 33657109, 33657182, 27486261, 27486262, 33657349, 27486285, 35695783, 18108374, 18108378, 18108377, 356968423, 35695855, 264631, 52644332, 264556, 18108385, 56526488, 87186516, 60432113, 264483
2005	87400884 (4008, 4010)	Novel Protein sim. GBank g1387350[emb]CA877951 - (Z47812) similar to ubiquitin carboxyl-terminal hydrolase; cDNA EST EMBL:033386 comes from this gene; cDNA EST EMBL:033985 comes from this gene; cDNA EST EMBL:033822 comes from this gene; cDNA EST EMBL:034547 comes from this gene...		ubiquitin	264488, 264906
2006	95351177 (4011, 4012)	Novel Protein sim. GBank g14106873[emb]CAA228131 - (AL035064) queuine tRNA-ribosyltransferase [Schizosaccharomyces pombe]	Contains protein domain (PF01702) - Queuine tRNA-ribosyltransferase	UNCLASSIFIED	56182575, 56181686, 22278995, 22278996, 22278997, 22278998, 22278999, 264259, 56182181, 60424269, 86714117, 35696052, 264906, 68712502, 264909, 264510, 60433356, 85658542, 265010, 265018, 265019, 284882, 284448, 284288, 284788, 28148627, 21906769, 29148784, 35695917, 60170615, 284891, 33857023, 66274820, 33657109, 55810784, 55811576, 35695855, 87186518, 60432113, 264563, 264482
2007	94326558 (4013, 4014)	Novel Protein sim. GBank g12662161[emb]BAA237121 - (AB007800) HH0452 cDNA clone for KIAA0440 has a 436-bp insertion at position 1711 of the sequence of KIAA0440. [Homo sapiens]		UNCLASSIFIED	264488, 263994, 35698052, 284508, 264905, 264906, 264907, 264908, 264909, 265007, 264910, 264592, 264595, 284756, 265011, 284780, 284762, 284764, 284766, 264685, 264767, 264768, 264769, 55811957, 35695917, 265020, 264691, 264693, 264628, 264629, 85274791, 35695855, 264631, 264632, 264634, 264635, 284837, 284838, 264839, 264585
2008	85084428 (4015, 4016)	Novel Protein sim. GBank g11550783[emb]CAA692571 - (Y07860) homeodomain protein [Mus musculus]	Contains protein domain (PF00046) - Homeobox domain	homeobox	264809, 264768, 35695855

2009	85748240 (4017, 4018)	Novel Protein sim. GBank gij3882305[idj]PAA34512.1] - (AB018335) KIAA0782 protein [Homo sapiens]		UNCLASSIFIED	22278999, 264299, 264910, 264591, 265017, 264681, 264683, 21906768, 264691, 33657182, 33857349, 264831, 87188518, 264404, 22279002, 264583
2010	85422458 (4018, 4020)	Novel Protein sim. GBank gij3262829[emb]CAB45735.1] - (AL080164) hypothetical protein [Homo sapiens]	Contains protein domain (PF00057) - eph Low-density lipoprotein receptor domain class A		52684507, 52684518, 52684365, 52684842, 18108397, 65274572, 22278994, 56994075, 35696286, 22278998, 22278997, 22278999, 264259, 26331822, 52684508, 29331824, 26331825, 29331828, 29331827, 29331828, 264511, 265007, 264512, 265008, 265009, 80432229, 80433358, 21906754, 52684317, 33108934, 52684298, 87188474, 87188559, 265017, 265018, 265019, 264881, 264885, 264687, 52684229, 264689, 21906765, 21906768, 21908787, 21908788, 35895917, 265020, 52684150, 264891, 264682, 33657023, 263967, 52684129, 35695763, 18108378, 35698423, 85274791, 35695855, 264831, 264634, 60431650, 264637, 264838, 52684332, 60170394, 18108385, 87188518, 22278002, 264564, 264565, 264568, 264567, 56182575, 56994075, 22278999, 264259, 29331824, 29331828, 29331827, 29331828, 35896052, 264906, 88712502, 265008, 265007, 265008, 265010, 265011, 265017, 265019, 264681, 284448, 264683, 264389, 264288, 264685, 264766, 264687, 21908785, 21906787, 21908768, 21908789, 265020, 265022, 264891, 33657023, 85274620, 33657109, 264629, 264557, 264559, 83373044, 87188518, 80432113, 22278002, 264259, 29331822, 29331824, 29331825, 29146498, 87188559, 265019, 264682, 264288, 264688, 21906764, 265020, 265022, 33657023, 264693, 33657109, 55811578, 264832, 264558, 58182323, 264839, 18108383, 18108384, 18108388, 22279000, 22279002, 264587
2012	87772137 (4023, 4024)	Novel Protein sim. GBank gij1086878 (U41020) - coded for by C. elegans cDNA YK10094.5; coded for by C. elegans helix-loop-helix-leucine zipper transcription factor [Caenorhabditis elegans]	Contains protein domain (PF00409) - Kinesin light chain repeat	UNCLASSIFIED	264259, 29331822, 29331824, 29331825, 29146498, 87188559, 265019, 264682, 264288, 264688, 21906764, 265020, 265022, 33657023, 264693, 33657109, 55811578, 264832, 264558, 58182323, 264839, 18108383, 18108384, 18108388, 22279000, 22279002, 264587
2013	94643642 (4025, 4026)	Novel Protein sim. GBank gij4507985[ref]NP_003427.1[pZNF1 - zinc finger protein 135 (clone pHZ-17)]	Contains protein domain (PF00098) - Zinc finger, C2H2 type	dna_mn_bind	18108398, 264808, 265007, 265010, 265018, 265019, 264689, 21906787, 265020, 264692
2014	87347840 (4027, 4028)	Novel Protein sim. GBank gij12720[sp]P209381[MYO_HETFR - MYELIN PO PROTEIN PRECURSOR]		UNCLASSIFIED	264488, 29331828, 264907, 264636, 264555, 264639, 264558

2015	88094822 (4029, 4030)	Novel Protein sim. GBank gijl812856jil[S22697 - extensin - Vohox carteri (fragment)]		UNCLASSIFIED	58182575, 35696288, 264259, 35696052, 264508, 264906, 264907, 264510, 264512, 87188474, 265010, 264681, 264286, 264889, 264826, 35696423, 35695855, 264638, 264563, 264564
2016	85286641 (4031, 4032)	Novel Protein sim. GBank gijl285046jil[S26413 - i-complex protein Tsp-10 - mouse]	struct	UNCLASSIFIED	264102, 264508, 264110, 265009, 33109954, 21906788, 265021, 33657109, 27488262, 263972, 18108374, 263976, 264555, 264564
2017	79464293 (4033, 4034)	Novel Protein sim. GBank gijl24735jilP18175jilNVO_FIG - INVOLUCRIN		UNCLASSIFIED	264885, 264636
2018	79637067 (4035, 4038)	Novel Protein sim. GBank gijl2143910jil[S68218 - phosphatase-1 glycogen-binding (GL)-chain - rat]	phosphatase		264107, 264110, 264112, 265017, 263976
2019	87787800 (4037, 4038)	Novel Protein sim. GBank gijl2072463 (U43200) - antifreeze glycoprotein AFGP polypeptide precursor [Boreogadus saida]	UNCLASSIFIED		264259, 264508, 264591, 265018, 264682, 264288, 264688, 22279002
2020	94674478 (4039, 4040)	Novel Protein sim. GBank gijl4216005 (AC008135) - putative Mcilin storage protein (globulin-like) [Arabidopsis thaliana]			56994075, 264593, 33109954, 21906754, 21908768, 33857023, 33857109, 27486281, 87168518
2021	86716918 (4041, 4042)	Novel Protein sim. GBank gijl585084jilP007803jilEFGM_RAT - ELONGATION FACTOR G, MITOCHONDRIAL PRECURSOR (MEF-G)			264757, 264767, 60170615, 16108385
2022	95285655 (4043, 4044)	Novel Protein sim. GBank gijl4216005 (AC008135) - putative Mcilin storage protein (globulin-like) [Arabidopsis thaliana]			
2023	87722976 (4045, 4046)	Novel Protein sim. GBank gijl5410230jilAAD42982, 1FA07334 - (AF07334) ubiquitin specific protease 3 [Homo sapiens]	Contains protein domain (PF00442) - ubiquitin hydrolases family 2	ubiquitin	16108394, 22278699, 264259, 264805, 264906, 264908, 264595, 264782, 264789, 264634, 264636, 87168516, 60432113, 22279000, 264482, 264585
2024	87886443 (4047, 4048)	Novel Protein sim. GBank gijl4755188jilAAD29055, 1JAC00701 - (JAC00701)			80433436, 265017, 264886, 264692, 264893, 264836
2025	87858863 (4049, 4050)	Novel Protein sim. GBank gijl4755188jilAAD29055, 1JAC00701 - (JAC00701)	Contains protein domain (PF00637) - 7-fold repeat in Clathrin and VPS	UNCLASSIFIED	22278997, 264508, 264906, 264909, 55812038, 265017, 265021, 265022, 80170615, 264556
2026	84122114 (4051, 4052)	Novel Protein sim. GBank gijl1655699jilCMA690321 - (Y07752) phenolphthalein-S [Vohox carteri]		UNCLASSIFIED	56994075, 60432048, 264508, 66712502, 264112, 60170831, 87168559, 264288, 264688, 264889, 21906766, 33657109, 18108370, 264638, 18108385, 60432113, 22279000, 22279002, 264564, 264566, 264587
2027	80249001 (4053, 4054)			UNCLASSIFIED	263976, 264634, 264486

2026	9469884 (4055, 4056)	Novel Protein sim. GBank g11330345 (U56735) - coded for by C. elegans cDNA yk34b1.5; coded for by C. elegans cDNA yk13n10.5; coded for by C. elegans cDNA yk46a8.5; coded for by C. elegans cDNA yk46d5.5; coded for by C. elegans cDNA yk43c2.5; coded for by C. elegans cDNA yk46e8.....	Contains protein domain (PF00632) - HECT-domain (ubiquitin-transferase).	ubiquitin	52644507, 52645158, 52646842, 56182575, 56894075, 35696286, 22276997, 22276998, 22276999, 60432049, 284259, 52845060, 29331822, 29331824, 66714117, 29331825, 29331826, 29331827, 29331828, 35696052, 264308, 264907, 264908, 29331830, 52844045, 56182435, 265006, 265009, 60432228, 33657402, 264595, 284757, 55812038, 21908754, 52846317, 52844296, 265010, 265011, 87168559, 265017, 265018, 265019, 264448, 18108354, 264288, 264389, 264786, 52844229, 21908765, 21908766, 21908767, 21908768, 21908769, 55611957, 35695917, 285020, 265021, 265022, 52844150, 33657023, 65274620, 33657109, 52645129, 16108368, 27486261, 27486262, 27486264, 27486265, 35695763, 264629, 55611576, 35696423, 35695855, 264635, 264636, 52844332, 264556, 83373044, 56526486, 22279000, 22279002, 264563, 265009, 264595, 65658542, 264555, 264556, 264557, 264558, 264559, 83373044
2029	85362032 (4057, 4058)	Novel Protein sim. GBank g13589940 (AF017368) - [acidogenital dysplasia protein 2 (Mus musculus)]	Contains protein domain (PF00621) - RhoGEF domain	UNCLASSIFIED	16108394, 56894075, 22276997, 22276999, 264259, 29331822, 29331824, 29331825, 66714117, 60432269, 29331826, 284106, 66712502, 264826, 265009, 265016, 265019, 264661, 264682, 264684, 264685, 56181562, 264686, 21908769, 265022, 264892, 264693, 264628, 18108370, 18108374, 264634, 264636, 264556, 16108365, 67166516, 22279002, 264565
2030	91213734 (4059, 4060)	Novel Protein sim. GBank g15630809p1AADA45023.1 (AC004690) similar to HUB1: similar to BAA24360 (P1Dg2769430) [Homo sapiens]	Contains protein domain (PF00096) - Zinc finger, C2H2 type	transcription factor	16108394, 56894075, 22276997, 22276999, 264259, 29331822, 29331824, 29331825, 66714117, 60432269, 29331826, 284106, 66712502, 264826, 265009, 265016, 265019, 264661, 264682, 264684, 264685, 56181562, 264686, 21908769, 265022, 264892, 264693, 264628, 18108370, 18108374, 264634, 264636, 264556, 16108365, 67166516, 22279002, 264565
2031	80245281 (4061, 4062)				264591, 55811957, 16108365, 264557, 264558, 18108382, 18108384
2032	91232607 (4063, 4064)	Novel Protein sim. GBank g15669491 (d013AAG3029.1) - (AB028000) KIAA1077 protein [Homo sapiens]	Contains protein domain (PF00864) - Sulfatase	hydrolase	65274572, 35696286, 29331824, 264908, 265009, 264593, 265018, 284288, 264686, 264769, 21908766, 21908767, 29148627, 284628, 35696423, 264634, 264558, 18108381, 60170394, 264558, 83373044, 18108385, 264482, 264484
2033	95000608 (4065, 4066)	Novel Protein sim. GBank g124948281p10B46861CAGT_RAT - ALPHA-N-ACETYLGLACTOSAMINIDE ALPHA-2-6-SIALYLTRANSFERASE (ST6GALNACIII) (STY)		synthase	56181562, 264628, 264632, 264555, 264556
2034	91232528 (4067, 4068)	Novel Protein sim. GBank g148269841ref1p_005147.1 (PFC01 - UNKNOWN RBD, or RNP domain)	Contains protein domain (PF00076) - RNA recognition motif (a.k.a. RRM, RBD, or RNP domain)	dna_rna_bind	22276998, 22276999, 264907, 29331830, 265008, 265018, 264681, 264682, 264684, 21908767, 21908768, 21908769, 33657108, 83373044, 56526486

2035	83553451 (4089, 4070)				264369, 264686, 265022, 56526486, 264587
2036	67115833 (4071, 4072)				29331827, 29331828, 264882, 264359,
					29148627, 60432113
2037	94324833 (4073, 4074)	Novel Protein sim. GBank g j2734061 (AF000195) - similar to oxysterol-binding proteins [Caenorhabditis elegans]		UNCLASSIFIED	65274572, 22278995, 22278996, 56994075, 35696286, 22278997, 22278998, 22278999, 264259, 29331824, 60432289, 29331828, 29331828, 35696052, 264907, 29331830, 66712502, 56182435, 265008, 265009, 80170831, 264594, 55812038, 33109954, 21906754, 87166559, 265017, 265016, 265019, 264782, 284389, 264288, 21906765, 21908767, 21906768, 21906769, 55811957, 35695917, 265020, 285021, 265022, 52644150, 33657023, 33657109, 33657182, 35695763, 35695855, 264832, 264834, 264836, 56182323, 63373044, 60432113, 22278000, 22279002, 264563
2038	95422384 (4075, 4076)	Novel Protein sim. GBank g j3806251 (emb CA807856) - (Z83785) predicted using GeneFinder, similar to RNA recognition motif (aka RRM, RBD, or RNP domain); cDNA EST EMBL:U01682 comes from this gene; cDNA EST EMBL:U75823 comes from this gene; cDNA EST EMBL:U27558 comes from this ge...	Contains protein domain (PF01412) - Putative GTP-ase activating protein for Arf	UNCLASSIFIED	22278995, 22278998, 56994075, 264259, 29331824, 35696052, 264905, 264906, 52644045, 265007, 265009, 87166559, 265017, 18108351, 284448, 264389, 264768, 264767, 264688, 18108358, 21908765, 21906769, 52644150, 33657023, 264692, 18108362, 33657109, 27486262, 18108370, 18108374, 18108379, 35896423, 85274791, 264832, 264836, 18108383, 83373044, 18108385, 87166516, 22278000, 22279002, 264563, 264564, 264568
2039	85514828 (4077, 4078)	Novel Protein sim. GBank g j224653 (daj gAA20813) - (AB002354) KIAA0356 [Homo sapiens]	Contains protein domain (PF00097) - Zinc finger, C3HC4 type (RING finger)	UNCLASSIFIED	22278997, 264259, 29331822, 264905, 264906, 264907, 264908, 264909, 264510, 265009, 264910, 264583, 264758, 285011, 265018, 284782, 284288, 264766, 264768, 264769, 21906766, 33657023, 264692, 264693, 33657109, 35696423, 264631, 264632, 264634, 264635, 264636, 264637, 264639, 67168518, 284488
2040	95308417 (4079, 4080)			UNCLASSIFIED	264592
2041	95071736 (4081, 4082)	Novel Protein sim. GBank g j2500825 (p 70700RPA2_MOUSE - DNA-DIRECTED RNA POLYMERASE I 135 KD POLYPEPTIDE (RNA POLYMERASE I SUBUNIT 2) (RPA135)		mapolymerase	264488, 22278998, 35696052, 264905, 264907, 264908, 264910, 265018, 264605, 265019, 18108351, 264768, 264769, 21906768, 265021, 265022, 264682, 33657109, 264828, 264629, 35696423, 35695855, 264637, 264838, 264583, 264584, 264585, 264567

2042	85307447 (4083, 4084)	Novel Protein sim. GBank gij4406590jg AA020040 - (AF131786) Similar to Era-VASP like protein [Homo sapiens]	Contains protein domain (PF00568) - WH1 domain	UNCLASSIFIED	60424179, 35696286, 264259, 29331828, 35696052, 29331828, 264508, 264509, 264907, 264908, 264510, 264511, 265008, 264910, 264591, 60433356, 264595, 265017, 265019, 264681, 264784, 264369, 264785, 264684, 264288, 264768, 264688, 52644229, 264769, 21808765, 35893917, 264535, 52644150, 264691, 264692, 18108365, 27486281, 27486282, 27486285, 18108374, 35896423, 85274791, 35895855, 264555, 264558, 60170394, 18108385, 264404, 22279000, 22279002, 264482, 264583, 264564, 264568
2043	84328076 (4085, 4086)	Novel Protein sim. GBank gij5052554jg AA038607 (AF145632) BcDNA, GH08032 [Drosophila melanogaster]	Contains protein domain (PF00122) - E1-E2 ATPase	transport	264488, 52644507, 52646365, 56994075, 22278997, 22278998, 20281171, 264259, 29331822, 29331824, 68714117, 29331828, 29331828, 33658970, 29148498, 264508, 264908, 52644045, 56182435, 265006, 33857402, 21906754, 52844296, 87188558, 265017, 265018, 265019, 264681, 264288, 264766, 264885, 264686, 21906766, 21906767, 21906768, 21906769, 265020, 265021, 80170815, 284691, 33857023, 284693, 65274620, 33657109, 33657182, 27488261, 27488262, 33657349, 35695763, 18108374, 55811576, 35695855, 18108380, 18108381, 80170394, 58182323, 284558, 83373044, 18108385, 56526486, 87168518, 60432113, 22279000, 264567
2044	87108927 (4087, 4088)	Novel Protein sim. GBank gij2248532 (U93872) - ORF 73, contains large complex repeat CR 73 [Kaposi's sarcoma-associated herpesvirus]		struct	264093, 29331827, 264905, 66712502, 264592, 284889, 21908785, 21908789, 265020, 264692, 264482, 264568
2045	78635532 (4089, 4090)	Novel Protein sim. GBank gij4406590jg AA020062 - (AF131852) Unknown [Homo sapiens]			264692
2046	87320849 (4091, 4092)	Novel Protein sim. GBank gij4101720 (AF-006466) - lymphocyte specific formin related protein [Mus musculus]			22278999, 29147620, 29331824, 29148498, 264508, 265007, 285008, 285019, 264805, 264681, 29148827, 28148628, 265021, 33857023, 18108365, 33857109, 33857182, 18108377, 264558, 264638, 264559, 18108388
2047	84578801 (4093, 4094)				264909
2048	84606378 (4095, 4096)	Novel Protein sim. GBank gij4589659jg BAA78850 1 - (AB023223) KIAA1008 protein [Homo sapiens]		UNCLASSIFIED	284488, 264259, 29331824, 29331828, 35696052, 264908, 264907, 264908, 264909, 264910, 264603, 264763, 21908787, 21908788, 264628, 264634, 264637, 22279002, 264564, 264585, 264586, 264567

2050	79633635 (4099, 4100)			UNCLASSIFIED	264693
2051	67780188 (4101, 4102)			UNCLASSIFIED	264488, 264259, 264509, 264906, 264907, 264789, 18108374, 35696423, 264583, 264586, 264488
2052	88086393 (4103, 4104)	Novel Protein sim. GBank gll4529889[gblpAAD21812.1]- (Af134728) G9A [Homo sapiens]	Contains protein domain (PF00656) - Kinase SET domain	UNCLASSIFIED	264488, 263994, 35696052, 264508, 264905, 264509, 264908, 264907, 264908, 264909, 264113, 264511, 285009, 264910, 60170831, 284592, 264758, 265010, 265011, 264605, 264780, 264682, 264784, 264389, 264786, 264688, 264788, 264789, 52644229, 264689, 35695917, 33657023, 33657109, 264628, 18108374, 35696423, 55811578, 35695955, 264630, 264631, 264632, 264634, 264635, 264636, 264556, 264638, 264639, 18108385, 56526486, 60432113, 264583, 264584, 264586, 264488, 264567, 264488, 263994, 35696052, 264508, 264905, 264509, 264906, 264907, 264908, 264909, 264113, 264511, 265009, 264910, 60170831, 264592, 264758, 265010, 265011, 264605, 264780, 264682, 264784, 264389, 264786, 264688, 52644229, 264689, 33657023, 33657109, 264628, 18108374, 35696423, 55811576, 35695955, 264630, 264631, 264632, 264634, 264635, 264636, 264556, 264638, 264639, 18108385, 56526486, 60432113, 264583, 264584, 264586, 264488, 264567
2053	87763078 (4105, 4106)	Novel Protein sim. GBank gll2995449[embCAA75113]- (Y14848) midline 1 protein [Mus musculus]		UNCLASSIFIED	22278998, 22278997, 264259, 26331822, 264102, 264508, 35695917, 263972, 264482
2054	95358937 (4107, 4108)	Novel Protein sim. GBank gll2876326[embCAB02090]- (Z79754) similar to C2 domain [Caenorhabditis elegans]	Contains protein domain (PF00168) - C2 domain		60424179, 264094, 264259, 26331825, 60424269, 264908, 60432229, 60433356, 87168559, 265019, 264780, 264288, 264688, 21908789, 33657023, 264693, 55810784, 55811578, 264635, 56182323, 60432113
2055	88259449 (4109, 4110)	Novel Protein sim. GBank gll5353748[gblpAAD42228.1]AF15913 - (AF159133) SIR2-like protein [Oryza sativa subsp. indica]		UNCLASSIFIED	264488, 29331828, 60432289, 29331828, 60433356, 265019, 264683, 264884, 265021, 33657109, 18108374, 264637, 18108385, 87188518, 60432113, 22279000, 264584

2056	86177396 (4111, 4112)	Novel Protein sim. GBank gl4628960refjnp_005042.1ipCARs - glutamine-tRNA synthetase	Contains protein domain (PT00749) - synthetase class I (E and O)	synthetase	264488, 52645156, 56182575, 22278994, 35696286, 58994075, 22278996, 22278998, 22278999, 60432049, 264259, 29331824, 60432289, 29331827, 29331828, 33656970, 264104, 264808, 264808, 265006, 265008, 60170831, 264591, 60432229, 60433438, 18108348, 21906734, 33657084, 52644298, 87168474, 265010, 87168559, 265017, 265018, 264760, 18108351, 264861, 264862, 264448, 264683, 264369, 264288, 264685, 294687, 264888, 264689, 21906765, 21906766, 21906767, 21906789, 55811957, 35695917, 265022, 33857023, 18108382, 33657109, 18108388, 33657182, 27486261, 27486264, 27486285, 33857349, 284628, 18108370, 264629, 18108374, 18108377, 18108378, 35696423, 55811576, 20261152, 264838, 264952, 18108385, 18108386, 87188518, 264482, 264585, 264588, 264587, 52646842, 52646365, 56182575, 35696286, 22278996, 22278997, 22278998, 264093, 52645080, 35696052, 29331828, 33656970, 265009, 52648317, 55811388, 52644296, 52644428, 21906768, 35695917, 265021, 60170815, 52644150, 33857109, 33657182, 27486261, 27486282, 35695763, 35696423, 35695855, 52644332 265007, 265008, 264591 29331825, 264882, 264686, 264691, 264693, 22279002
2057	87877605 (4113, 4114)	Novel Protein sim. GBank gl12288501p066401AMYH_YEAST - GLUCOAMYLASE S1S2 PRECURSOR (GLUCAN 1,4-ALPHA- GLUCOSIDASE) (1,4-ALPHA-D-GLUCAN GLUCOHYDROLASE)		UNCLASSIFIED	
2058	86276896 (4115, 4116)	Novel Protein sim. GBank gl1197141sp13983EXTN_TOBAC - EXTENSIN PRECURSOR (CELL WALL HYDROXYPROLINE-RICH GLYCOPROTEIN)		UNCLASSIFIED	
2059	78866684 (4117, 4118)	Novel Protein sim. GBank gl12811122 (U67318) - NaDC-2 [Xenopus laevis]		UNCLASSIFIED	
2060	83050800 (4119, 4120)	Novel Protein sim. GBank gl12811122 (U67318) - NaDC-2 [Xenopus laevis]		UNCLASSIFIED	

2081	95362204 (4121, 4122)	Novel Protein sim. GBank gil2496947sp Q08289 YQO9_CAEEL - HYPOTHETICAL 141.2 KD PROTEIN EEDD8.9 IN CHROMOSOME II	Contains protein domain (PF00069) Eukaryotic protein kinase domain	kinase	22278997, 22278999, 264259, 29331822, 29331824, 29331826, 29331827, 35696052, 29331828, 264906, 66712502, 29331830, 264808, 264909, 264112, 264511, 265007, 265008, 264910, 264591, 33657402, 21906754, 85658542, 265017, 265019, 264448, 264683, 264288, 264684, 264369, 264686, 264687, 18108358, 264889, 21806765, 21806767, 21806768, 21806769, 265020, 265022, 264691, 33657023, 33657109, 20281149, 18108379, 35695855, 264634, 264558, 264557, 264558, 18108382, 264559, 83373044, 18108384, 58528488, 60432113
2082	87028440 (4123, 4124)	Novel Protein sim. GBank gil4502081 ref NP_001139.1 pANK2 - ankryn 2, neuronal	Contains protein domain (PF00023) Ank repeat	struct	264905, 264628, 264907, 264629, 264908, 264909, 18108374, 263978, 35695855, 264512, 264635, 60431850, 264638, 264760, 264563, 18108351, 264782, 264585, 264784, 264487, 264768
2083	87801272 (4125, 4126)	Novel Protein sim. GBank gil4589582 db BAI76803.1 - (AB023178) KIAA0959 protein [Homo sapiens]	Contains protein domain (PF00617) RasGEF domain	oncogene	22278994, 22278998, 264258, 29331827, 264906, 264909, 52844045, 264886, 21906767, 55811957, 264692, 18108365, 263972, 55811576, 18108384, 22279002, 264482, 264563, 264564, 264484
2084	85317253 (4127, 4128)	Novel Protein sim. GBank gil1754515 db BAI13413.1 - (D87515) aminopeptidase-B [Rattus norvegicus]		hydrolase	264488, 52845365, 56894075, 35696286, 22278997, 22278998, 264259, 29331828, 60432288, 29331827, 29331828, 35696052, 264509, 265007, 265008, 60432229, 60433438, 21906754, 265010, 265011, 87168558, 265017, 265018, 264781, 18108351, 264682, 264368, 264288, 52844229, 21906765, 21908787, 21908788, 35685917, 33857109, 18108366, 18108374, 35898423, 35895855, 52844332, 264559, 60432113, 22279000, 22279002, 264566, 264488
2085	85082238 (4129, 4130)	Novel Protein sim. GBank gil2507144 sp Q04205 TENS_CHICK - TENSIN		kinase	264568, 18108394, 56182181, 60432289, 29331828, 264905, 264908, 264808, 60431735, 60433358, 55811386, 85658542, 265018, 55811150, 264881, 264788, 264692, 60431526, 263974, 55810764, 35695855, 264831, 264634, 264635, 60431850, 264557, 83373044, 18108388, 22279000, 22279002, 55182575, 284259, 264906, 264764, 264288, 58182323, 264567
2086	85793402 (4131, 4132)	Novel Protein sim. GBank gil160171 (M58295) - circumsporozoite protein [Plasmodium yoelii]	Contains protein domain (PF00096) Zinc finger, C2H2 type	UNCLASSIFIED	

2067	85303892 (4133, 4134)				35696288, 22278997, 22278998, 60433049, 264259, 60432289, 80433438, 264682, 264448, 264369, 264288, 18108355, 21908765, 21906768, 265022, 33657109, 35698423, 35695855, 284558, 264404, 264563, 264488
2068	84344754 (4135, 4136)			UNCLASSIFIED	264687
2069	94319177 (4137, 4138)	Novel Protein sim. GBank g j3152662 (AF084604) - KEG3 protein [Homo sapiens]	Contains protein domain (PF00023) - Ank repeat	transcriptador	60424178, 56182575, 22278995, 22278996, 56994075, 204259, 29331822, 29331824, 29331825, 35696052, 29331828, 33658970, 264509, 264905, 58182435, 265009, 60433358, 87188559, 265017, 265018, 264604, 265019, 264448, 264764, 264766, 21906785, 21906787, 21906768, 21906769, 265020, 265021, 33657023, 33657109, 283976, 264555, 264557, 56182323, 83373044, 87166518, 60432113, 22279000, 22279002
2070	85791380 (4139, 4140)	Novel Protein sim. GBank g j5712131gb ACD47379.1 AF12049 - (AF120499) DEM1 protein [Homo sapiens]		UNCLASSIFIED	35695917, 264905, 264628, 264908, 264638
2071	86946116 (4141, 4142)	Novel Protein sim. GBank g j3551531db BAA330161 - (AB017437) avena [Gallus gallus]	Contains protein domain (PF00558) - WH1 domain		18108398, 265006, 265007, 265008, 265009, 264594, 265010, 265011, 18108351, 18108354, 18108364, 18108385, 18108388, 264634, 18108381, 18108385, 18108388, 18108391
2072	91716429 (4143, 4144)		Contains protein domain (PF00184) - Neurohypophysial hormones, C-terminal Domain		22278995, 35696286, 22278997, 22278998, 22278999, 264490, 60433049, 264259, 29331822, 29331824, 29331826, 35696052, 265008, 33657402, 21906754, 265011, 265018, 18108351, 264682, 264369, 21906765, 21906766, 21906767, 21906768, 21906769, 35695917, 265020, 265021, 264680, 264682, 35696423, 264555, 264556, 284558, 22279000
2073	27825684 (4145, 4146)	Novel Protein sim. GBank g j1504028 db BAA132121 - (D88878) similar to C.elegans protein (Z37093) [Homo sapiens]		UNCLASSIFIED	264556
2074	94324787 (4147, 4148)	Novel Protein sim. GBank g j240317 db BAA14937.1 - (AB020721) KIAA0914 protein [Homo sapiens]			29331822, 264809, 264511, 265009, 264594, 264595, 265010, 265011, 265017, 265018, 265019, 264448, 264683, 265020, 265021, 18108370, 264632, 63373044, 264587

2075	84314866 (4149, 4150)	Novel Protein sim. GBank gl 5138930 gb AA0382.11 - (AF093880) transcription factor. t1b [Homo sapiens]			18108394, 22278994, 22278996, 35696288, 22278998, 22278999, 264259, 29331822, 29331825, 29331827, 35698052, 29331828, 264905, 264907, 264908, 264510, 265007, 264910, 265009, 33657084, 264760, 264448, 264288, 264768, 264767, 264689, 21908786, 21608767, 21908789, 265021, 265022, 60170615, 33657023, 27486262, 27486265, 35698423, 35698555, 264631, 264834, 264639, 87168518, 22279002, 264563, 264486, 18108391
2076	87564118 (4151, 4152)				264259, 29331828, 264508, 264908, 264510, 265007, 265011, 264286, 264637, 18108385
2077	11389877 (4153, 4154)			UNCLASSIFIED	264592
2078	87539384 (4155, 4156)	Novel Protein sim. GBank gl 4220590 gb BA745791 - (D87908) nuclear protein np25 [Mus musculus]	Contains protein domain (PF00626) - PHD-finger		29331825, 265017, 265018, 264288, 285020, 265021, 264634, 58526488
2079	88085916 (4157, 4158)	Novel Protein sim. GBank gl 4240255 gb BA74906.1 - (AB020690) KIAA0883 protein [Homo sapiens]	Contains protein domain (PF00098) - Zinc finger, CCHC class	UNCLASSIFIED	264768, 22278997, 265021, 264690, 264259, 264892, 29331822, 264693, 29331824, 29331828, 264508, 264509, 264907, 264628, 20281089, 264909, 265007, 285009, 264832, 264638, 264591, 264592, 264839, 264768, 264759, 33109954, 284804, 285018, 265019, 22279002, 264563, 264564, 264448, 284684, 264567, 264885
2080	84136689 (4159, 4160)	Novel Protein sim. GBank gl 2408021 emb CAB18219.1 - (Z98162) putative vacuolar protein [Schizosaccharomyces pombe]	Contains protein domain (PF01363) - FYVE zinc finger	UNCLASSIFIED	56182575, 264082, 29331824, 29331826, 29331830, 285017, 285018, 265020, 83373044
2081	94847186 (4161, 4162)	Novel Protein sim. GBank gl 5524734 gb AA044360.1 - (AF16635) ST7 protein [Homo sapiens]	Contains protein domain (PF00431) - CUB domain	epb	22278996, 22278997, 22278999, 60432049, 29331822, 29331824, 33657402, 65856542, 265011, 265018, 265019, 21908787, 21908788, 264693, 18108385, 22279000, 22279002
2082	87828629 (4163, 4164)	Novel Protein sim. GBank gl 3890558 emb CAA942341 - (Z70271) predicted using GeneFinder: similar to collagen, CDNA EST yk308e7.3 comes from this gene: CDNA EST yk308e7.5 comes from this gene: CDNA EST yk385a8.3 comes from this gene: CDNA EST yk385a8.5 comes from this gene [Caeno...]		collagen	264807, 285019
2083	84141000 (4165, 4166)	Novel Protein sim. GBank gl 2352427 - (AF004161) - peroxisomal Ca-dependent solute carrier [Oryctolagus cuniculus]	Contains protein domain (PF00153) - Mitochondrial carrier proteins	transport	56182575, 22278996, 22278997, 284259, 29331822, 29331825, 264509, 264112, 265009, 264593, 21908784, 265018, 285019, 264448, 284288, 264885, 21908788, 285020, 265022, 264691, 18108370, 85274791, 264631, 264555, 264583
2084	95199298 (4167, 4168)	Novel Protein sim. GBank gl 28636 sp P39193 ALU6 - IIII ALU SUBFAMILY SP WARNING ENTRY IIII		UNCLASSIFIED	285018, 264763, 264663, 264691

2085	84989476 (4169, 4170)	Novel Protein sim. GBank g11655699jemb[CA469032] - (Y0752), phosphoin-S [Volvox carter]	UNCLASSIFIED	56182575, 80432289, 264908, 56182435, 87166474, 264783, 264369, 264888, 264693, 18108370, 58182323
2086	91234404 (4171, 4172)	Novel Protein sim. GBank g13875032jemb[CA468936] - (Z49125) similarly to Trichostonyx colubriformis 11 kd secretory protein (Swiss Prot accession number P21937); cDNA EST EMBL.D33349 comes from this gene; cDNA EST EMBL.D37844 comes from this gene; cDNA EST EMBL.D36149 come...	UNCLASSIFIED	35696286, 264259, 35686052, 264906, 264807, 264908, 264908, 264910, 264759, 264604, 264762, 264768, 264769, 35695917, 263978, 35696423, 35695855, 264632, 264634, 264637, 264838, 264639, 58182323, 18108385, 264482, 264488
2087	21436337 (4173, 4174)	Novel Protein sim. GBank g10880930jemb[CA416334.1] - (AL021481) similar to Phosphoglucosylase and phosphomannomutase phosphoserine; cDNA EST EMBL.D36188 comes from this gene; cDNA EST EMBL.D70697 comes from this gene; cDNA EST YK3739.5 comes from this gene; cDNA EST EMBL.T0080...	UNCLASSIFIED	264488, 22278994, 35696288, 22278996, 29331827, 35696052, 35657402, 21906754, 33108954, 87166474, 265017, 265018, 265019, 264448, 264883, 264388, 264685, 264687, 264688, 21908765, 21906768, 21906767, 21906768, 21906769, 265020, 265021, 265022, 264692, 35657023, 35657109, 35657182, 27486261, 27486262, 35657349, 27486285, 35696423, 35695855, 83373044, 87166518, 22278000, 264587
2088	94111527 (4175, 4176)	Novel Protein sim. GBank g10880930jemb[CA416334.1] - (AL021481) similar to Phosphoglucosylase and phosphomannomutase phosphoserine; cDNA EST EMBL.D36188 comes from this gene; cDNA EST EMBL.D70697 comes from this gene; cDNA EST YK3739.5 comes from this gene; cDNA EST EMBL.T0080...	UNCLASSIFIED	18108392, 264488, 52644507, 18108394, 18108397, 52648842, 18108398, 56182575, 22278994, 22278995, 35696288, 22278998, 56994075, 22278997, 22278998, 22278999, 264091, 264092, 264093, 264094, 60432049, 264259, 26331822, 20281099, 29331824, 29331825, 29331826, 29331827, 29331828, 35696052, 35696052, 29146498, 29146499, 264102, 264108, 264107, 264109, 264508, 264905, 264509, 264908, 264907, 264908, 66712502, 264628, 52644045, 264509, 56182435, 264110, 264112, 264510, 264511, 265008, 264512, 265007, 265008, 264910, 265009, 60170631, 264592, 264593, 80433356, 35657402, 80433438, 264595, 55812038, 264758, 21906754, 35857084, 55811386, 52644286, 265010, 265011, 87168559, 265017, 265018, 265019, 264760, 264781, 55811150, 264762, 18108351, 264682, 264448, 264783, 264784, 264883, 264369, 18108354, 264288, 264685, 264766, 264888, 264687, 264768, 52644229, 264688, 18108358, 56181562, 264789, 18108359, 264889, 21906785, 21906766, 21906767, 21906768, 29148627, 21908768, 55811957, 29148629, 29148784, 35885917, 265020, 265021, 265022, 60170615, 264690
2089	95422801 (4177, 4178)	Novel Protein sim. GBank g14758118jemb[IP_004623.1]PCAP3 - Death associated protein 3	cadherin	

2090	88222470 (4179, 4180)				22278995, 22278998, 22278999, 264259, 29331826, 35686052, 264910, 33657402, 60433438, 53109954, 8718474, 87168559, 265018, 265019, 264681, 264684, 264686, 264687, 264688, 264689, 21908765, 21906766, 21906767, 21908768, 35695917, 265022, 60170615, 33657023, 35696423, 35695855, 264952, 18108387, 22279000
2091	95309161 (4181, 4182)	Novel Protein sim. GBank g14580971g9pAAD24571.1(A-F12108 - (A-F121081) cAMP inducible 2 protein [Mus musculus]		UNCLASSIFIED	263994, 264905, 264906, 264511, 264512, 265008, 264910, 55611388, 284288, 264768, 56181562, 21906763, 21906768, 21906769, 265022, 264628, 264563, 264567
2092	88223605 (4183, 4184)			homeobox	22278997, 22278999, 66712502, 87168559, 264683, 265021, 264488
2093	87406073 (4185, 4186)	Novel Protein sim. GBank g12352427 (A-F004161) - peroxisomal Ca-dependent solute carrier [Oryctolagus cuniculus]	Contains protein domain (PF00153) - Mitochondrial carrier proteins	transport	264905, 264906, 264907, 264908, 264510, 265006, 265007, 265009, 264910, 264598, 21906754, 87168474, 265011, 284603, 265018, 265019, 264760, 264766, 264768, 264769, 21906767, 21906768, 21906769, 265021, 264680, 33657023, 264693, 264628, 264634, 264636, 264837, 264557, 56182323, 264584
2094	91230828 (4187, 4188)	Novel Protein sim. GBank g14929551g9pAAD34036.1(A-F15179 - (A-F151799) CGI-40 protein [Homo sapiens]		MHC	35696286, 265017, 265018, 265019, 18108388
2095	95351526 (4189, 4190)	Novel Protein sim. GBank g113632381g9pA57284 - spermelid perinuclear RNA-binding protein Spnr - mouse	Contains protein domain (PF00035) - Double-stranded RNA binding motif	dna_rne_bind	35696286, 52644045, 265006, 265007, 265008, 87168559, 18108351, 21908789, 29148784, 265020, 33657023, 27486282, 18108374, 18108388
2096	84118760 (4191, 4192)	Novel Protein sim. GBank g13834423 (A-F070688) - Cytoplasmic dynein intermediate chain isoform DIC1a [Drosophila melanogaster]	Contains protein domain (PF00400) - WVD domain, G-beta repeat	ATPase_associated	264488, 264489, 85274572, 56182575, 22278996, 22278997, 22278999, 284259, 60432289, 28331826, 35696052, 264107, 264508, 264509, 264905, 264906, 264907, 284508, 52644045, 264909, 264510, 264511, 264512, 265008, 264910, 265009, 264582, 60433356, 60433438, 264758, 264596, 55812038, 21908754, 284601, 284602, 264605, 264762, 264681, 16108351, 264764, 264683, 264288, 264687, 264768, 264769, 264689, 21906765, 21906766, 21906767, 35695917, 285020, 265022, 52644150, 264691, 264692, 33657023, 264693, 27486281, 35695763, 264628, 264629, 35696423, 35695855, 264831, 264632, 264634, 264635, 284555, 284837, 263981, 264638, 264639, 284563, 284483, 264565, 264566, 264489, 264567

2007	95322772 (4193, 4194)	Novel Protein sim. GBank g15174501 e1fNP_008051.1 p1YF1 - zinc finger protein, subfamily 1A, 1 (kairos)	Contains protein domain (PF00086) - Zinc finger, C2H2 type	Transcript factor	65274572, 264511, 265010, 264600, 265017, 284448, 264288, 265021, 60170815, 264692, 33657109, 18108370, 264636, 264483 56994075, 264258, 264286, 265020, 264563
2098	67760340 (4195, 4196)	Novel Protein sim. GBank g14756209 e1fNP_004081.1 pDUSP - dual specificity phosphatase 3 (vaccinia virus phosphatase V1H-related)	Contains protein domain (PF00782) - Dual specificity phosphatase, catalytic domain	phosphatase	65274572, 264905, 65274444, 264691, 264636, 264555
2099	95412927 (4197, 4198)	Novel Protein sim. GBank g12695653 (AF026954) - pyruvate dehydrogenase phosphatase regulatory subunit precursor, PDP1 (Bos laurus)			
2100	95332656 (4199, 4200)	Novel Protein sim. GBank g13381188 e1fNP_004081.1 pDUSP - dual specificity phosphatase 3 (vaccinia virus phosphatase V1H-related) EMBL: C08179 comes from this gene: cDNA EST EMBL: C08337 comes from this gene: cDNA EST EMBL: C08829 comes from this gene: cDNA EST YK291b4.5 comes from this gene: cDNA EST YK4....	Contains protein domain (PF00025) - ADP-ribosylation factor family	nucl_repl	56182575, 22278995, 22278996, 22278997, 22278998, 60432049, 264259, 29331622, 29331824, 29331825, 29331827, 29331828, 29146498, 264909, 265006, 265009, 264910, 264591, 60432229, 60433356, 33657402, 264758, 21906754, 85658542, 87166474, 265017, 265016, 265019, 264601, 18108351, 264762, 264448, 264369, 264288, 18108355, 264886, 21906765, 21906767, 21906768, 21906769, 265020, 265021, 33657023, 18108374, 35686423, 264558, 83373044, 87168518, 60432113, 22279000, 22279002 264091, 29331824, 264105, 265007, 265010, 18108380
2101	67762604 (4201, 4202)	Novel Protein sim. GBank g14589468 e1fNP_004081.1 pDUSP - dual specificity phosphatase 3 (vaccinia virus phosphatase V1H-related) EMBL: C08179 comes from this gene: cDNA EST EMBL: C08337 comes from this gene: cDNA EST EMBL: C08829 comes from this gene: cDNA EST YK291b4.5 comes from this gene: cDNA EST YK4....		UNCLASSIFIED	264408, 264489, 35696286, 264259, 35696052, 264508, 264905, 264907, 264908, 264909, 264511, 264512, 264591, 264593, 60433356, 264758, 264601, 264605, 264780, 18108351, 264448, 264784, 264288, 264767, 264768, 21906769, 35695917, 18108374, 264834, 264555, 264559, 264563, 264482, 264486
2102	87770461 (4203, 4204)	Novel Protein sim. GBank g13874149 e1fNP_004081.1 pDUSP - dual specificity phosphatase 3 (vaccinia virus phosphatase V1H-related) EMBL: C08179 comes from this gene: cDNA EST EMBL: C08337 comes from this gene: cDNA EST EMBL: C08829 comes from this gene: cDNA EST YK291b4.5 comes from this gene: cDNA EST YK4....		UNCLASSIFIED	264408, 264489, 35696286, 264259, 35696052, 264508, 264905, 264907, 264908, 264909, 264511, 264512, 264591, 264593, 60433356, 264758, 264601, 264605, 264780, 18108351, 264448, 264784, 264288, 264767, 264768, 21906769, 35695917, 18108374, 264834, 264555, 264559, 264563, 264482, 264486
2103	95413576 (4205, 4206)	Novel Protein sim. GBank g14240159 e1fNP_004081.1 pDUSP - dual specificity phosphatase 3 (vaccinia virus phosphatase V1H-related) EMBL: C08179 comes from this gene: cDNA EST EMBL: C08337 comes from this gene: cDNA EST EMBL: C08829 comes from this gene: cDNA EST YK291b4.5 comes from this gene: cDNA EST YK4....	Contains protein domain (PF01530) - Zinc finger, C2HC type	Transcript factor	65274572, 56994075, 22278998, 264258, 29331824, 29331825, 35696052, 28331828, 66712502, 265009, 80170631, 264595, 33109954, 85658542, 87166536, 265017, 265019, 264448, 21906765, 21906768, 265022, 33657023, 27488262, 33657348, 35695763, 60431528, 18108374, 55811576, 56182323, 18108387, 67168516, 60432113, 264584
2104	85776161 (4207, 4208)			UNCLASSIFIED	264582, 264604, 22279000

2105	84640080 (4209, 4210)	Novel Protein sim. GBank gll1707032 (JB045) - coded for by C. elegans CDNA yk1365.3; coded for by C. elegans CDNA yk1268.3; coded for by C. elegans CDNA CEMSE18F; coded for by C. elegans CDNA yk1260.1.3; coded for by C. elegans CDNA yk65h6.3; coded for by C. elegans CDNA yk65h6....		UNCLASSIFIED	264488, 56162575, 22278994, 56994075, 22278998, 22278997, 22278998, 22278999, 60432049, 264259, 29331822, 29331824, 29331826, 60432289, 29331827, 29331828, 294508, 294905, 294509, 294907, 29331830, 52644045, 264510, 264511, 265007, 264512, 265009, 80170831, 60432229, 33657402, 60433356, 264595, 60433438, 264758, 33657084, 67166474, 265010, 87166559, 285017, 265018, 265019, 264762, 18108351, 264684, 18108354, 264288, 264688, 52644229, 18108359, 21906785, 21906788, 21906787, 21906788, 21906789, 35685817, 265020, 265022, 60170615, 52644150, 264692, 33657023, 284693, 33657109, 60431528, 18108374, 65274781, 35685855, 264635, 60170394, 264639, 264558, 18108385, 18108387, 56526486, 87168518, 60432113, 264564, 264566, 264567, 265006, 265019
2106	83365475 (4211, 4212)	Novel Protein sim. GBank gll3881524lemb[CAS3883] - (Z70038) ZK1067.4 [Caenorhabditis elegans]		UNCLASSIFIED	264906, 264639
2107	79822662 (4213, 4214)	Novel Protein sim. GBank gll3176689 (AC003671) - Contains similarity to ubiquitin carboxy-terminal hydrolase 14 gblZ35927 from S. cerevisiae [Arabidopsis thaliana]		UNCLASSIFIED	264905, 264906, 264907, 284908, 264909, 284738, 265011, 264600, 264601, 264784, 264766, 284787, 284788, 284789, 284693, 264629, 35695855, 284632, 264634, 264635, 264638, 264639, 83373044, 264486, 18108348, 264769, 18108370, 18108374, 284555, 264556, 284557, 264558
2108	80478719 (4217, 4218)			UNCLASSIFIED	284564
2110	87729075 (4219, 4220)	Novel Protein sim. GBank gll481043pjl[S37671] - ba12 protein - human		UNCLASSIFIED	264766, 35685917, 264630, 264567, 284486, 284508, 284908, 284581, 284882, 22275002
2111	87618419 (4221, 4222)	Novel Protein sim. GBank gll2143639pjl[S56542] - calmodulin-binding protein - rat	Contains protein domain (PF00059) - struct	UNCLASSIFIED	265006
2112	87283783 (4223, 4224)	Novel Protein sim. GBank gll426823gbl[AAD20459] - (AF100960) proteadherin [Rattus norvegicus]	Eukaryotic protein kinase domain		66741117, 29331826, 29331827, 60433438, 55812038, 265017, 265019, 264689, 21906789, 55811957, 265020, 265021, 33657109, 60170394, 284558
2113	78941388 (4225, 4226)	Novel Protein sim. GBank gll426823gbl[AAD20459] - (AB014585) KIAA0885 protein [Homo sapiens]	Contains protein domain (PF00028) - cadherin		65274572, 264689, 264691, 264692, 60432113
2114	87689342 (4227, 4228)			UNCLASSIFIED	
2115	80983785 (4229, 4230)	Novel Protein sim. GBank gll4757690pjl[NP_004328, flpC6OR - chromosome 8 open reading frame 1		UNCLASSIFIED	

2116	88258387 (4231, 4232)	Novel Protein sim. GBank gij224552 (U83872) - ORF 73, contains large complex repeat CR 73 [Kaposi's sarcoma-associated herpesvirus]		struct	35696288, 22278999, 56182181, 28331824, 28331825, 28331827, 35696052, 264907, 56182435, 285008, 284591, 55812038, 55811388, 87168558, 264288, 284368, 21908789, 28148629, 33857023, 35695763, 55811578, 35698423, 18108385
2117	87788904 (4233, 4234)	Novel Protein sim. GBank gij2330021 (AF019250) - kinesin-related protein; KRP, Costal2 [Drosophila melanogaster]		struct	28331824, 264511, 285008, 33109954, 285017, 285018, 284288, 284889, 285020, 284892, 58526488, 264482
2118	87078894 (4235, 4236)	Novel Protein sim. GBank gij1079307 (J155573) - nuclear pore complex glycoprotein p62 - African clawed frog		glycoprotein	284259, 284905, 284907, 284908, 284510, 284511, 285009, 284910, 285010, 284602, 284288, 284768, 284893, 283987, 283972, 284638, 284559
2119	86999317 (4237, 4238)	Novel Protein sim. GBank gij4321407 (J157481) - (AF047690) ATP-binding cassette protein M-ABC1 [Homo sapiens]		UNCLASSIFIED	284893, 18108385
2120	87788395 (4239, 4240)	Novel Protein sim. GBank gij4885527 (refNP_005480.1) (NSP3 - novel SH2-containing protein 3	Contains protein domain (PF00017) - eph Src homology domain 2	eph	284091, 284259, 28331828, 28331828, 285017, 284604, 284288, 284885, 285020, 284691, 18108370, 55810764, 284555, 284838, 80432113
2121	80021375 (4241, 4242)	Novel Protein sim. GBank gij4757728 (refNP_004886.1) (PAGTA - angiotensin vasopressin receptor AII/AVP-like		UNCLASSIFIED	284601, 284766, 263978
2122	91230931 (4243, 4244)	Novel Protein sim. GBank gij4829551 (J157481) (AF15179) CGI-40 protein [Homo sapiens]			18108394, 56182575, 22278997, 29331822, 29331824, 28331825, 28331828, 28331828, 284907, 56182435, 285007, 284910, 285010, 285018, 284888, 285020, 55811578, 284555, 284637, 18108382, 83373044, 18108383, 18108384, 58528488, 284565, 284567, 18108388, 284757, 285011, 18108385, 284691, 284634, 18108385
2123	86787898 (4245, 4246)	Novel Protein sim. GBank gij2224551 (J157481) (AF047690) KIAA0305 [Homo sapiens]	Contains protein domain (PF01363) - struct FYVE zinc finger	struct	29331822, 284906, 284907, 284591, 284639, 284591, 284634, 18108385
2124	83005951 (4247, 4248)	Novel Protein sim. GBank gij5689455 (J157481) (AF047690) KIAA1059 protein [Homo sapiens]	Contains protein domain (PF00801) - PKD domain	transport	284591, 284634, 18108385
2125	95354041 (4249, 4250)	Novel Protein sim. GBank gij728831 (J157481) (AF047690) KIAA1059 protein [Homo sapiens]		UNCLASSIFIED	284591, 284634, 18108385
2126	95354041 (4249, 4250)	Novel Protein sim. GBank gij728831 (J157481) (AF047690) KIAA1059 protein [Homo sapiens]		UNCLASSIFIED	284591, 284634, 18108385
2127	95354041 (4249, 4250)	Novel Protein sim. GBank gij728831 (J157481) (AF047690) KIAA1059 protein [Homo sapiens]		UNCLASSIFIED	284591, 284634, 18108385
2128	95084231 (4251, 4252)	Novel Protein sim. GBank gij4539264 (J157481) (AF047690) conserved hypothetical protein [Schistosoma pombe]		UNCLASSIFIED	284591, 284634, 18108385

2127	81118652 (4253, 4254)	Novel Protein sim. GBank gik686435gb/AUC31315.1(AE14323) - (AE14323) apoptosis related protein AFR-2 [Homo sapiens]			35696286, 29331828, 35696052, 284506, 284509, 284905, 284908, 284907, 284908, 284909, 284510, 285008, 284511, 284512, 285007, 285009, 284910, 284758, 285011, 284600, 284601, 284604, 284782, 284783, 284786, 284687, 284768, 284769, 284689, 35695917, 284690, 284691, 284692, 284693, 284629, 18108374, 35695855, 284632, 284634, 284635, 284637, 284638, 284639, 18108385, 284503, 284564, 284567, 56182575, 35698288, 56182181, 29331824, 60432289, 35696052, 284805, 284807, 66712502, 284908, 284909, 284510, 284512, 285009, 284910, 284591, 55812038, 285018, 284764, 284288, 284369, 284687, 284788, 55811957, 284692, 18108368, 284628, 284632, 284634, 284635, 284637, 56182323, 284639, 18108384, 18108386, 284563, 284567
2128	87414262 (4255, 4256)				
2129	95102089 (4257, 4258)			UNCLASSIFIED	66714117, 284828, 284595, 55812038, 55811150, 55811957, 284693, 18108374, 283978, 65274791, 18108381, 63373044, 22279000
2130	95417144 (4259, 4260)	Novel Protein sim. GBank gik2648235 (AE001012) - conserved hypothetical protein [Archaeoglobus fulgidus]		UNCLASSIFIED	283981
2131	65723085 (4261, 4262)	Novel Protein sim. GBank gik1068886 (U41276) - Similar to potassium channel protein [Caenorhabditis elegans]	Contains protein domain (PF00805) - Pentapeptide repeats (8 copies)	potassium_channel	35696052, 284909, 284768, 35695917
2132	95361098 (4263, 4264)	Novel Protein sim. GBank gik5689373(djBAA82973.1) - (AB028944) KIAA1021 protein [Homo sapiens]	Contains protein domain (PF00122) - E1-E2 ATPase	ATPase_associated	284486, 22278999, 284259, 29331827, 28331628, 35696052, 284509, 284905, 284908, 284907, 284908, 284909, 284510, 285008, 284511, 285007, 284910, 284591, 284592, 284595, 284758, 21908754, 33109954, 87168474, 285011, 284600, 284601, 284605, 285019, 284780, 18108351, 284681, 284782, 284764, 284288, 284684, 284786, 284686, 284687, 284768, 284769, 284688, 21908769, 284690, 52844150, 284691, 284693, 18108370, 284628, 284629, 18108372, 18108374, 35696423, 35695855, 284631, 284634, 284635, 284636, 284555, 284637, 18108380, 284639, 284558, 56182323, 56528486, 284564, 284565, 284568, 284567

2133	95351539 (4265, 4268)	Novel Protein sim. GBank gij4220489 (AC006069) - hypothetical protein (Arabidopsis thaliana)		UNCLASSIFIED	60424178, 52646365, 52646842, 56694075, 35686286, 22278997, 22278998, 60432049, 58182181, 66714117, 60424269, 29331828, 29331828, 35686052, 264905, 264906, 264907, 66712502, 29331830, 56182435, 265006, 264512, 265008, 60431735, 60433358, 33657402, 53612038, 33109954, 21908754, 55811388, 265010, 264603, 265017, 265018, 265019, 55811150, 18108351, 264682, 264369, 284288, 52644229, 56181562, 21906765, 21906766, 21906767, 21906768, 21906769, 55811957, 35685917, 265020, 265021, 60170615, 33657023, 33657109, 60431528, 18108374, 35686423, 65274791, 35685855, 264634, 60431650, 18108380, 56182323, 83373044, 18108385, 18108387, 60432113, 22279002, 264563, 264566
2134	95412687 (4267, 4268)	Novel Protein sim. GBank gij3875351emb (CAB09415) - (Z96047) D13.6 [Caenorhabditis elegans]			56181688, 35686286, 21906754, 55811388, 265011, 265017, 18108351, 264765, 264768, 264688, 21906768, 35685917, 265020, 33657023, 264628, 35685855, 264632, 264555, 264556, 264557, 264558, 18108382, 22279002
2135	88079613 (4269, 4270)	Novel Protein sim. GBank gij5689559djb (AA83063.1) - (AB029034) KIAA1111 protein [Homo sapiens]	Contains protein domain (PF00628) - PHD-finger	UNCLASSIFIED	22278999, 29331828, 35686052, 264906, 264908, 264910, 265009, 264591, 264758, 52646317, 265011, 87188559, 264601, 18108351, 264448, 264883, 264884, 264689, 18108359, 264691, 33657023, 264692, 35685763, 264829, 35685855, 264631, 264635, 264638, 264637, 56182323, 264639, 22279002, 264584
2136	84346478 (4271, 4272)	Novel Protein sim. GBank gij2662167dbj (BA23715) - (AB007903) KIAA0443 [Homo sapiens]		UNCLASSIFIED	264569, 264909, 33109954, 264763, 21908768, 60170394, 18108385, 264563
2137	87637716 (4273, 4274)	Novel Protein sim. GBank gij4984110emb (CAB43262.1) - (AL050090) hypothetical protein [Homo sapiens]		UNCLASSIFIED	264239, 29331828, 35686052, 264909, 285008, 265017, 265018, 18108351, 264288, 21908768, 33657023, 33657109, 264628, 18108374, 35685855, 264634, 264555, 264556, 264557, 264558, 264559
2138	87385446 (4275, 4276)	Novel Protein sim. GBank gij5174779gb (A040896.1) - (U67804) 50 kDa protein [Caulobacter crescentus]		ATPase, associated	264905, 264910, 264991, 55812038, 55811388, 85685842, 264760, 18108351, 18108359, 55811957, 265020, 265021, 33657023, 18108364, 55811576, 83373044, 18108385, 56828486, 264482
2139	94643882 (4277, 4278)	Novel Protein sim. GBank gij3850821emb (CAA77135) - (Y18350) U2 snRNP auxiliary factor, large subunit [Nicotiana glauca]			

2140	87646651 (4279, 4280)	Novel Protein sim. GBank gll4417293jpbAA0204181 - (ACO07018) unknown protein [Arabidopsis thaliana]	UNCLASSIFIED	264468, 264259, 29331824, 264104, 264109, 264509, 265006, 264759, 265018, 264448, 264288, 21908768, 55911957, 265021, 33657023, 27486265, 35686423, 264836, 264556, 264557, 264559, 264566
2141	79823988 (4281, 4282)		UNCLASSIFIED	265020, 264693
2142	80041222 (4283, 4284)		UNCLASSIFIED	263978
2143	94140051 (4285, 4286)	Novel Protein sim. GBank gll2135766jpbSS3362 - mucin 5AC (clone JER47) - human (fragment)	UNCLASSIFIED	22276997, 29331827, 264907, 265020, 60432113
2144	94320114 (4287, 4288)	Novel Protein sim. GBank gll2076463 (U43200) - antifreeze glycoprotein AFGP polypeptide precursor [Boreogadus saida]	UNCLASSIFIED	65274572, 264259, 29331824, 29331827, 264908, 264909, 264591, 265011, 87188559, 264600, 265019, 264288, 264768, 21908765, 21906767, 55911576, 35686423, 85274791, 22279002
2145	20564305 (4289, 4290)		UNCLASSIFIED	263978
2146	87010515 (4291, 4292)	Novel Protein sim. GBank gll1255871 (U53341) - short region of weak similarity to bovine membrane receptor p63 (PIR:S28503) [Caenorhabditis elegans]	UNCLASSIFIED	264909, 60433356, 264688
2147	80432911 (4293, 4294)	Novel Protein sim. GBank gll5060389jpbCAA18718.11 - (AL022603) putative protein [Arabidopsis thaliana]	UNCLASSIFIED	264907, 264768, 264769, 16108335
2148	80048811 (4295, 4296)	Novel Protein sim. GBank gll728837jpbP39194/ALU7_HUMAN - III ALU SUBFAMILY SQ WARNING ENTRY III	UNCLASSIFIED	264593
2149	87362022 (4297, 4298)	Novel Protein sim. GBank gll119883jpbP20693jCE2_MOUSE - LOW AFFINITY IMMUNOGLOBULIN EPSILON FC RECEPTOR (LYMPHOCYTE IGE RECEPTOR) (FC-EPSILON-RII) (CD23)	Contains protein domain (PF000059) - Lectin C-type domain	29331824, 29331826, 35699052, 264756, 87188474, 265018, 52644150, 33657109
2150	84140059 (4299, 4300)	Novel Protein sim. GBank gll5420387jpbCAB46679.11 - (AJ243459) proteophosphoglycan [Leishmania major]	UNCLASSIFIED	22276998, 29331822, 29331824, 29331828, 264764, 264769, 21908768, 264488
2151	95535241 (4301, 4302)	Novel Protein sim. GBank gll56694407jpbBAA82887.11 - (AB028956) KIAA1035 protein [Homo sapiens]	UNCLASSIFIED	22276998, 29331822, 29331824, 29331828, 264764, 264769, 21908768, 264488, 22276998, 56994075, 22276998, 60432049, 264259, 29331822, 29331824, 29331826, 35696052, 29331828, 264508, 264511, 80433356, 264758, 264598, 33109954, 60174639, 265010, 265011, 87168559, 265017, 265018, 265019, 264448, 264288, 264689, 21908765, 21908768, 21908768, 265020, 60170615, 33657109, 33657182, 33657349, 16108370, 264635, 264557, 60170394, 16108365, 87188518, 22279000
2152	79321640 (4303, 4304)	Novel Protein sim. GBank gll3452473 (AF084205) - serine/threonine protein kinase TAO1 [Rattus norvegicus]	kinase	16106397, 16106398, 265007, 264591, 265011, 16108351, 16108366, 16108374, 16108388

2153	66313371 (4305, 4308)	Novel Protein sim. GBank g14758704[re]NP_004218.1[PMASL - MTH-amplified sequences with leucine-rich tandem repeats 1]	Contains protein domain (PF00560) - glycoprotein Leucine Rich Repeat	UNCLASSIFIED	264488, 263894, 52646842, 22278996, 22278998, 22278999, 264259, 29331822, 35698052, 264508, 264509, 264905, 264906, 264907, 264908, 264909, 56182435, 264510, 264511, 264512, 264758, 87188474, 87186559, 265017, 265019, 264780, 264288, 264389, 264786, 264887, 264769, 52844229, 21908768, 21908769, 35695917, 33657023, 33657109, 35695955, 264631, 264832, 264635, 264838, 264839, 18108385, 264483, 264564, 264468
2154	87408034 (4307, 4308)	Novel Protein sim. GBank g1223150[pr]11209265U - chorion protein B11 [Bombyx mori]		UNCLASSIFIED	56984075, 264094, 265009, 265019, 264288, 21908767, 35695917
2155	87424072 (4309, 4310)			UNCLASSIFIED	18108392, 18108398, 22278998, 264239, 29331824, 265008, 265010, 265011, 265017, 265019, 264288, 264886, 265020, 264693, 264628, 58182323
2156	64285205 (4311, 4312)	Novel Protein sim. GBank g13970966 (AC004974) - spa-1-like, similar to AF028504 (P1Dg2555163) [Homo sapiens]			265007, 264664
2157	87316344 (4313, 4314)	Novel Protein sim. GBank g11076211[pr]11550755 - hypothetical protein VSP-3 - Chlamydomonas reinhardtii		UNCLASSIFIED	264591
2158	80083729 (4317, 4318)	Novel Protein sim. GBank g14650844[db]BA7027.11 - (AB026190) Kelch motif containing protein [Homo sapiens]	Contains protein domain (PF00651) - dna_ma_bind		264598
2160	18263674 (4319, 4320)	Novel Protein sim. GBank g12879925[db]BA248261 - (AB007897) KIAA0437 [Homo sapiens]	BTB/POZ domain		29331822, 264112, 285009, 264681, 33657023, 264634
2181	87739131 (4321, 4322)	Novel Protein sim. GBank g11504006[db]BA132021 - (D88958) simlanto human ZFY protein. [Homo sapiens]		UNCLASSIFIED	264834
2162	84319528 (4323, 4324)			UNCLASSIFIED	265008
2163	95417158 (4325, 4328)	Novel Protein sim. GBank g13676537[emb]CA982701 - (Z73974) cDNA EST YK29115.3 comes from this gene; cDNA EST YK29115.5 comes from this gene [Caenorhabditis elegans]		UNCLASSIFIED	65274572, 264508, 264805, 264806, 264807, 264808, 52644045, 264809, 265007, 264910, 264591, 264592, 264593, 55812038, 264598, 264758, 265011, 264600, 264762, 264763, 264683, 264764, 264288, 264788, 264888, 264768, 264769, 264688, 265020, 264691, 264628, 264829, 283978, 264832, 264834, 264557, 264838, 264639, 18108385, 284563, 264566, 264587
2164	80569456 (4327, 4328)		Contains protein domain (PF01006) - collagen Hepatitis C virus non-structural protein NS4a		58182575, 22278998, 264093, 264683, 33657023, 65274820, 80432113
					264603, 264637, 264555

2165	04328169 (4328, 4330)	Novel Protein sim. GBank g11086794 (U41107) - No definition line found [Caenorhabditis elegans]	UNCLASSIFIED	56994075, 22278996, 22278997, 22278998, 264258, 29331822, 29331824, 29331825, 29331826, 29331827, 29331828, 264908, 29331830, 56182435, 285009, 21906754, 33657084, 265011, 265018, 264448, 264288, 264389, 21906765, 21906768, 21906789, 265020, 265021, 264691, 264692, 33657023, 65274620, 35695855, 264556, 60170394, 83373044, 80432113, 22278992, 264567, 52645156, 22278994, 22278998, 66714117, 28331828, 52644045, 265018, 265019, 264369, 21906765, 21906787, 21906768, 21906769, 265021, 265022, 264693, 27466262, 35695783, 18108376, 58526468, 87169518, 264567
2166	07818834 (4331, 4332)	Novel Protein sim. GBank g12706522jcm[CA75816] - (Y15895) ubiquitin activating enzyme [Drosophila melanogaster]	ubiquitin	56182575, 35696286, 29331824, 29331828, 29146488, 56182435, 285008, 285009, 264592, 264593, 33657402, 33108954, 265011, 265017, 265018, 18108351, 264389, 21906764, 21906765, 21906768, 28148627, 21908769, 52644150, 33657109, 35696423, 18108381, 18108384, 18108385, 60432113, 264567
2167	67716864 (4333, 4334)	Novel Protein sim. GBank g12224713[db]BA20840] - (AB002384) KIAA0386 [Homo sapiens]	UNCLASSIFIED	66714117, 29331827, 264907, 264511, 264591, 265018, 264764, 264883, 284768, 264788, 284566
2168	86899334 (4335, 4336)	Novel Protein sim. GBank g14321407[gb]AAD15748] - (AF047690) ATP-binding cassette protein M-ABC1 [Homo sapiens]	Contains protein domain (PF00664) - transport ABC transporter transmembrane region.	264529, 264555, 264559
2169	87868937 (4337, 4338)	Novel Protein sim. GBank g15108521[gb]AAD39741.1[AF10536] K-CI cotransporter KCCL4 [Homo sapiens]	UNCLASSIFIED	65274572, 58182575, 22278997, 22278998, 284259, 29331825, 264509, 264908, 56182435, 60433438, 55812038, 264598, 55811386, 265018, 264782, 264763, 264448, 264764, 264684, 264288, 264766, 264685, 58181592, 264688, 55811857, 265020, 264535, 264691, 33657109, 60431528, 18108374, 35696423, 55811576, 65274791, 264634, 264638, 264558, 87168518, 60432113, 284564
2170	94141033 (4339, 4340)	Novel Protein sim. GBank g15108521[gb]AAD39741.1[AF10536] K-CI cotransporter KCCL4 [Homo sapiens]	UNCLASSIFIED	264389, 285020, 284558
2171	80194050 (4341, 4342)	Novel Protein sim. GBank g14308681[gb]AAD15478] - (AC008930) R33423.1 [Homo sapiens]	UNCLASSIFIED	264389
2172	85452460 (4343, 4344)	Novel Protein sim. GBank g12493778[sp]C09458[YO35_CAEEL - PUTATIVE CUTICLE COLLAGEN C09G5.5]	UNCLASSIFIED	264908, 35695855, 284555, 264557
2173	87038740 (4345, 4346)	Novel Protein sim. GBank g12493778[sp]C09458[YO35_CAEEL - PUTATIVE CUTICLE COLLAGEN C09G5.5]	UNCLASSIFIED	
2174	95003286 (4347, 4348)	Novel Protein sim. GBank g12493778[sp]C09458[YO35_CAEEL - PUTATIVE CUTICLE COLLAGEN C09G5.5]	UNCLASSIFIED	

2175	94325650 (4348, 4350)	Novel Protein sim. GBank g 1263287 (U47655) - fibrin-3 [Araenus diadematus]		UNCLASSIFIED	264486, 35696286, 20281089, 29331628, 60432289, 35696052, 264109, 264508, 264509, 264905, 264906, 264907, 264908, 264909, 264510, 264511, 265006, 265007, 264512, 265008, 264910, 264591, 264594, 264595, 264596, 264758, 55612038, 265011, 264600, 264603, 264760, 264762, 264448, 264764, 264288, 264766, 264888, 264687, 21906768, 55811957, 35695917, 265020, 265022, 264691, 264692, 33657023, 264693, 264628, 264629, 55811576, 35696423, 65274791, 35695855, 264631, 264632, 264634, 264635, 264636, 264637, 264638, 264639, 264558, 18108365, 60432113, 264563, 264564, 264565, 264566, 264486, 264567
2176	88223392 (4351, 4352)	Novel Protein sim. GBank g 126837sp P36194 ALU7_HUMAN - III ALU SUBFAMILY SQ WARNING ENTRY III	Contains protein domain (PF00805) - oncogene Pentapeptide repeats (6 copies)	oncogene	52644507, 52646642, 22278994, 35686286, 22278996, 22278998, 29331826, 29331827, 35686052, 29331628, 33656970, 29331830, 264910, 33657402, 264758, 52644298, 87168559, 265016, 264689, 21906765, 21906767, 21906769, 35695917, 52644150, 264690, 33657023, 33657109, 52645129, 33657162, 27486261, 27486262, 33657349, 18108376, 18108377, 35695855, 87166516, 60432113, 264404, 22278900, 264486
2177	94128942 (4353, 4354)	Novel Protein sim. GBank g 15454072 p NP_006416.1 pSLU7 - step II splicing factor SLU7		Kinase	16106592, 22278997, 22278999, 264093, 33657402, 265019, 264446, 264766, 264689, 21906767, 21906768, 21906769, 265021, 33657023, 18108370, 18108374, 60432113, 22278902
2178	87601557 (4355, 4356)	Novel Protein sim. GBank g 473407 (U08215) - NST-1 [Mus musculus]	Contains protein domain (PF00012) - eph Hsp70 protein	eph	264488, 22278996, 22278999, 29331824, 29331825, 29331828, 29331827, 52644296, 87168474, 18108370, 35695855, 22278902, 60424269, 264760, 264628, 264632
2179	87316275 (4357, 4358)			UNCLASSIFIED	

2180	95351387 (4359, 4360)	Novel Protein sim. GBank g 3122317 sp P0848 KIMB_DICD1 - MYOSIN HEAVY CHAIN KINASE B (MHCK B)	Contains protein domain (PF00400) - WD domain, G-beta repeat	Kinase	52644507, 22278994, 35696286, 22278997, 22278999, 284259, 52845080, 29331822, 29331824, 29331825, 29331826, 29331828, 33658970, 284508, 284509, 284508, 284508, 29331630, 284509, 284510, 285008, 285007, 33657402, 35612038, 21906754, 87166474, 87188559, 285017, 285018, 285019, 284763, 284682, 284683, 284684, 284268, 284686, 21906765, 21906768, 21906769, 285020, 285021, 285022, 52844150, 33657023, 33657109, 27486265, 33657349, 18108374, 35698423, 35698425, 283981, 80170394, 18108385, 58528488, 87166518, 60432113, 22278000, 284482, 284568, 284587, 284488 29331827, 284389, 18108378, 284564
2181	85784830 (4361, 4362)	Novel Protein sim. GBank g 3024889 sp Q15542 T2D4_HUMAN - TRANSCRIPTION INITIATION FACTOR TFIID 100 KD SUBUNIT (TAFII-100) (TAFII100)		Kinase	
2182	87637731 (4363, 4364)	Novel Protein sim. GBank g 15420367 emb CAB4879.1 - (AJ234359) proteophosphoglycan [Leishmania major]		UNCLASSIFIED	22278986, 22278997, 22278999, 264259, 29331822, 58182435, 284112, 264784, 284288, 21906767, 21908768, 21906769, 33657109, 18108378, 80170394, 22279000, 22279002
2183	85480849 (4365, 4366)	Novel Protein sim. GBank g 1873406 gb ACT7482.1 - (U17129) unknown [Rhodococcus erythropolis]		inf	264760
2184	87760890 (4367, 4368)	Novel Protein sim. GBank g 3114713 (AF061346) - Edp1 protein [Mus musculus]			29331822, 29331825, 29331826, 56182435, 285011, 284685, 284686, 21906768, 18108370, 284829, 284631, 284636, 284557 29331824, 284907, 88712502, 264757, 285019, 284288, 284692, 56526488
2185	87828463 (4369, 4370)	Novel Protein sim. GBank g 15106958 gb AAC039806.1 (AF113815) FH1/FH2 domain-containing protein FHOS (Homo sapiens)		ATPase-associated	264259, 29331822, 29331824, 29331826, 58182435, 284592, 5812038, 284780, 284786, 5811957, 33657023, 5811578, 56182323, 284563
2186	87738227 (4371, 4372)	Novel Protein sim. GBank g 12864625 emb CAA18972 - (AL021811) putative protein [Arabidopsis thaliana]		ATPase-associated	33696052, 284905, 284906, 284907, 284908, 284510, 284511, 285008, 284910, 284756, 285018, 284782, 284881, 284768, 284769, 33696917, 284692, 33696423, 284631, 284635, 284637, 18108388, 284568, 284488 18108398, 585994075, 284259, 29331824, 29331825, 68714117, 29331827, 284908, 29331830, 285018, 285020, 285021, 58182323, 284558, 22279000, 22279002
2187	87386173 (4373, 4374)			UNCLASSIFIED	
2188	87777708 (4375, 4378)	Novel Protein sim. GBank g 15107816 gb AAC0129.1 (AF149413) contains similarity to histone deacetylases. Pfam PF00650. Score=13.3, E=5e-10, N=1 [Arabidopsis thaliana]		histone	
2189	85693573 (4377, 4378)	Novel Protein sim. GBank g 15452357 (AF075724) - unknown [Legionella pneumophila]	Contains protein domain (PF01599) - C-methyltransferase		22278998, 284259, 29331828, 21908754, 284389, 284288, 283987

2190	87639197 (4378, 4380)	Novel Protein sim. GBank gij3257siep28315fRNI_RAT - RIBONUCLEASE INHIBITOR		nucleaseinhb	22278996, 22278999, 29331822, 28331824, 29331828, 265008, 264910, 60170831, 35812038, 52844298, 265010, 265018, 264684, 264888, 56181562, 21908769, 35695917, 265022, 60170394, 22278000, 29331825, 29331828, 29331830, 264510, 264511, 264910, 284583, 264584, 264558, 264559
2191	85198828 (4381, 4382)	Novel Protein sim. GBank gij5327002jemhCAB46272.11 - (Y18503) XAP-5-like protein [Homo sapiens]			
2192	11126316 (4383, 4384)	Novel Protein sim. GBank gij46260jsep34400jmi10 CAEEL - MIG-10 PROTEIN PH domain	Contains protein domain (PF00169) - PH domain		264558
2193	94140073 (4385, 4388)	Novel Protein sim. GBank gij5420389jemhCAB46680.11 - (AJ243460) prodeophosphoglycan [Leishmania major]		UNCLASSIFIED	56181686, 28331825, 28331827, 264508, 264509, 265008, 284592, 80432229, 264288, 264684, 264788, 35895917, 33657023, 60431802, 80431528, 55810764, 55811576, 85274791, 35695855, 80431850, 56182323, 60432113
2194	21418714 (4387, 4388)	Novel Protein sim. GBank gij2773341 (AF040954) - putative protein phosphatase 1 nuclear targeting subunit [Rattus norvegicus]			284592
2195	88083023 (4389, 4390)	Novel Protein sim. GBank gij2832783jemhCAA15685.11 - (AL008191) /prediction=(method: /prediction=(method: /match=(desc: /match=(desc: /motif=(desc: [Drosophila melanogaster]		UNCLASSIFIED	22278996, 22278999, 35688052, 285006, 21908754, 265017, 35695917, 285021, 265022, 35685855
2196	95091831 (4391, 4392)	Novel Protein sim. GBank gij5262407jemhCAB45699.11 - (AL080076) hypothetical protein [Homo sapiens]		collagen	56182575, 35696286, 22278997, 22278998, 264259, 28331822, 88714117, 60432289, 29331827, 35696052, 29331828, 284508, 52844045, 56182435, 264510, 265007, 265008, 285009, 60433438, 55812038, 265010, 285011, 284448, 264288, 284688, 284887, 52644229, 21908785, 21908788, 21908787, 35895917, 285022, 264691, 33857023, 264693, 18108370, 18108376, 35696423, 55811578, 85274791, 35695855, 264838, 58182323, 18108385
2197	85073813 (4393, 4394)	Novel Protein sim. GBank gij4929587j8bjiAD34044.11AF15180 - (AF151807) CGI-49 protein [Homo sapiens]			284788, 264789, 21908765, 21908768, 21908767, 29148827, 55811957, 35696286, 265020, 22278998, 265021, 284259, 33857023, 284893, 29331824, 35696052, 29331828, 18108370, 35695855, 264113, 265008, 264910, 80432229, 56182323, 33857402, 264758, 83373044, 21908754, 265016, 265019, 22279002, 284482, 284448, 284585, 284288, 284369
2198	88060914 (4395, 4396)	Novel Protein sim. GBank gij3548787 (AC005622) - R30953.1 [Homo sapiens]		UNCLASSIFIED	

2199	88054355 (4397, 4398)	Novel Protein sim. GBank gij2739372 (AC002505) - hypothetical protein [Arabidopsis thaliana]			284105, 284110, 284112, 284688, 55811957, 33657023, 284892, 283967, 20281071,
2200	87405385 (4398, 4400)	Novel Protein sim. GBank gij3043634dbjBA254811 - (AB011127) KIAA0555 protein [Homo sapiens]		struct	28331824, 284763, 284768 56526486
2201	94316872 (4401, 4402)	Novel Protein sim. GBank gij3913470spj057314DHEX_ANAPL - PUTATIVE STEROID DEHYDROGENASE SPW2	Contains protein domain (PF00106) - short chain dehydrogenase	dehydrogenase	28331824, 35696052, 284805, 264907, 33657402, 55811386, 265017, 285018, 265019, 264288, 21906768, 35695917, 285020, 285022, 33657023, 33657109, 27486261, 18108370, 35696423, 35695855, 264555, 284558, 83373044, 87188518, 80432113
2202	87672385 (4403, 4404)	Novel Protein sim. GBank gij362665fembCAB45787.11 - (AL080186) hypothetical protein [Homo sapiens]		UNCLASSIFIED	284489, 284259, 28331824, 60432289, 35696052, 284805, 264909, 264592, 265017, 265018, 265019, 18108351, 284782, 284448, 264389, 284288, 264768, 21908765, 21906768, 264690, 264691, 264692, 33657108, 264634, 264636, 264555, 264639, 284558, 284559, 83373044, 18108385, 264404, 22279002, 264482
2203	87761832 (4405, 4406)	Novel Protein sim. GBank gij1172845spjP48829R825_RABIT - RAS-RELATED PROTEIN RAB-25	Contains protein domain (PF00071) - Ras family	glycoprotein	52646385, 56994075, 284259, 29331822, 29331828, 29331827, 29331828, 284910, 265010, 265011, 87186559, 265018, 265019, 284605, 264288, 21906769, 35695917, 33857023, 284892, 33657109, 35695783, 18108376, 264638, 22278000, 264586, 264567
2204	88086871 (4407, 4408)	Novel Protein sim. GBank gij21036spjP2934gic813_RAT GUANINE NUCLEOTIDE-BINDING PROTEIN GGT, ALPHA- 3 SUBUNIT (GUSTOUCIN ALPHA-3 CHAIN)	Contains protein domain (PF00503) - G-protein alpha subunit	UNCLASSIFIED	
2205	94147589 (4409, 4410)	Novel Protein sim. GBank gij4589480dijBA78788.11 - (AB023141) KIAA0824 protein [Homo sapiens]	Contains protein domain (PF00086) - Zinc finger, C2H2 type	dna_mn_bind	18108394, 18108397, 56182575, 60432049, 264259, 20331822, 29331824, 29331825, 29331828, 29331827, 284808, 285007, 265008, 265009, 60432229, 285010, 265011, 265018, 284683, 284288, 264369, 264686, 21906768, 21908768, 21908789, 264890, 264691, 264683, 18108388, 55611576, 65274791, 264634, 18108381, 18108384, 60432113, 22279002, 284563, 284568 264581
2206	20820008 (4411, 4412)			UNCLASSIFIED	284581
2207	87767870 (4413, 4414)	Novel Protein sim. GBank gij4557753feijNP_000372.1pMID1 - midline 1 protein	Contains protein domain (PF00822) - SPRY domain		29331822, 58182181, 29331827, 35696052, 52644045, 265008, 265019, 58181582, 55811957, 265021, 33657023, 35695763, 35695855, 60170394, 80432113, 284588 264806, 265019, 18108351, 21906788 264112, 265009, 264891, 18108385, 18108374, 264634, 20281168
2208	88100830 (4415, 4416)			struct	
2209	87800420 (4417, 4418)	Novel Protein sim. GBank gij3986748 (AF105228) - lutein [Bos taurus]			

2210	57152407 (4418, 4420)	Novel Protein sim. GBank glt728837 sp P39194 ALU7_HUMAN - III ALU SUBFAMILY SQ WARNING ENTRY III		kinase	264603
2211	87341720 (4421, 4422)	Novel Protein sim. GBank glt728837 sp P39194 ALU7_HUMAN - III ALU SUBFAMILY SQ WARNING ENTRY III		oncogene	264685, 264686, 18108365, 22279002, 264482
2212	81223924 (4423, 4424)	Novel Protein sim. GBank glt376027 emb CAA092141 - (AJ010475) RNA helicase [Arabidopsis thaliana]	Contains protein domain (PF00270) - DEAD/DEAH box helicase	helicase	22276995, 22276897, 22276898, 264092, 264094, 29331822, 66714117, 29331828, 29331828, 264807, 52644045, 265009, 60170831, 21908754, 87186559, 265017, 265019, 18108351, 264883, 18108354, 264369, 264768, 264687, 52644229, 21908785, 21908788, 21908787, 21908788, 265021, 33657109, 18108370, 18108374, 264838, 58182323, 18108384, 18108387, 87188518, 264585
2213	81218308 (4425, 4428)	Novel Protein sim. GBank glt5420387 emb CAB46878.11 - (AJ243459) proteophosphoglycan [Leishmania major]			56182575, 22276998, 22276997, 35696052, 264905, 88712502, 264908, 264628, 56182435, 264112, 285008, 60431735, 60433438, 21906754, 265010, 265011, 265017, 265018, 265019, 18108351, 264765, 21908765, 21906768, 21908769, 265020, 265021, 264693, 284829, 263974, 263978, 18108378, 55811576, 264558, 264637, 264558, 83373044, 22278002, 264482, 284483

2214	95361453 (4427, 4428)	Novel Protein sim. GBank gll4504325/ie NP_000173.1 pHADH - hydroxyacyl-Coenzyme A dehydrogenase/3-ketoacyl-Coenzyme A thiolase/enoyl-Coenzyme A hydratase (infunctional protein), alpha su	Contains protein domain (PF00729) - 3-hydroxyacyl-CoA dehydrogenase	dehydrogenase	264468, 52644507, 18108394, 56182575, 22276994, 22276895, 35698268, 56984075, 22276997, 22276998, 22276899, 264490, 60432049, 264259, 52645080, 29331822, 29147620, 29331824, 66714117, 29331825, 60432289, 29331826, 29331827, 35696052, 29331828, 20281100, 264508, 264807, 66712502, 264806, 29331830, 52644045, 56182435, 264510, 265008, 264511, 264512, 265007, 265008, 265009, 60170831, 264593, 60433356, 60433436, 33109954, 33657064, 52644286, 67168474, 265010, 265011, 87168558, 264601, 265017, 265018, 265019, 16106351, 264448, 264682, 264763, 264288, 264687, 52644229, 264689, 21906765, 21906766, 21906767, 21906768, 21906769, 35695917, 265020, 265021, 265022, 264532, 60170615, 264690, 52644150, 264691, 33657023, 264692, 18108364, 33657109, 33657182, 27486262, 27486264, 27486265, 35695763, 18106370, 264629, 60431528, 16106374, 16108376, 55810764, 35686423, 35685655, 264634, 264636, 52644332, 264638, 264558, 80170394, 16106381, 56182323, 63373044, 18108385, 18108367, 16108388, 56526486, 67168518, 60432113, 22279002, 264482, 264564, 264565, 264566, 264909, 265006, 264555, 264556, 67168518
2215	95418206 (4429, 4430)	Novel Protein sim. GBank gll1947160 (AF000286) - weak similarity to colleagues, glycine- and proline-rich [Caenorhabditis elegans]			
2216	87614046 (4431, 4432)	Novel Protein sim. GBank gll1572802 (U70654) - similar to Enterococcus faecalis TRAB (GI:386268) [Caenorhabditis elegans]		UNCLASSIFIED	264693
2217	80569404 (4433, 4434)	Novel Protein sim. GBank gll5031707/ie NP_005503.1 pGARP - glycoprotein A repetitions predominant	Contains protein domain (PF00560) - Leucine Rich Repeat	glycoprotein	264288, 33657108, 264556
2218	85518254 (4435, 4436)	Novel Protein sim. GBank gll3678636/ie NP_005503.1 pGARP - glycoprotein A repetitions predominant (Z49126) similar to cAMP-dependent protein kinase; cDNA EST EMBL:700718 comes from this gene; cDNA EST yk46508.3 comes from this gene; cDNA EST yk46508.5 comes from this gene; cDNA EST yk4924.3 comes from this gene; cDNA EST y...	Contains protein domain (PF00069) - Eukaryotic protein kinase domain		35696423, 264563
2219	87614046 (4437, 4438)	Novel Protein sim. GBank gll1572802 (U70654) - similar to Enterococcus faecalis TRAB (GI:386268) [Caenorhabditis elegans]	Contains protein domain (PF01993) - Tri8 family		264662, 264683, 264686, 264689, 264693, 18106370, 16108376

2220	95354165 (4439, 4440)	Novel Protein sim. GBank gij4507201 [hefjNP_003145.1tpSTAT - stathefin			264488, 18108394, 18108395, 35696286, 264259, 264097, 60432289, 264508, 264905, 264906, 264907, 29331830, 264908, 264909, 264510, 264511, 285007, 264512, 264910, 265009, 264593, 264594, 60433358, 264595, 55812038, 264758, 85658542, 265010, 264601, 264603, 265019, 264605, 264760, 264762, 264448, 264764, 264369, 264766, 18108357, 264768, 264687, 18108358, 264769, 55811957, 264690, 264691, 33657023, 264692, 18108362, 18108368, 264628, 264629, 18108374, 263978, 264634, 264635, 264636, 264637, 264638, 18108385, 264483, 264565, 264486, 264567
2221	89080827 (4441, 4442)	Novel Protein sim. GBank gij3549154 (AC005625) - R27328_1 [Homo sapiens]			
2222	84425892 (4443, 4444)			UNCLASSIFIED	264908, 265020, 35695855
2223	95091849 (4445, 4446)			UNCLASSIFIED	265010, 264665, 264690, 264693, 264626, 263974, 263976, 55811576, 264555, 264638, 83373044, 264483
2224	87386515 (4447, 4448)	Novel Protein sim. GBank gij876005 [embjCAA84788] - (Z35719) cDNA EST EMBL: D67419 comes from this gene: cDNA EST EMBL: C13853 comes from this gene: cDNA EST yk234a7.3 comes from this gene: cDNA EST yk234a7.5 comes from this gene: cDNA ES...	Contains protein domain (PF01956) - Domain of unknown function	UNCLASSIFIED	264259, 264509, 56162435, 265006, 265008, 265009, 264757, 21906754, 18108351, 264693, 16108374, 18108385
2225	85749484 (4449, 4450)	Novel Protein sim. GBank gij1255847 (U53338) - C05E11.1 gene product [Caenorhabditis elegans]		transport	22278994, 22278995, 22278999, 52844045, 264600, 265019, 21908765, 21908769, 264259, 29331822, 29331824, 29331825, 29331827, 264506, 264908, 265007, 264691, 264834, 264486
2226	86978953 (4451, 4452)	Novel Protein sim. GBank gij4826524 [embjCAB42852.1] - (AL049848) hypothetical protein [Homo sapiens]			22278999, 265008, 265009, 18108354, 29148628, 29146764, 27486281, 18108374, 264637, 18108364
2227	87721135 (4453, 4454)			UNCLASSIFIED	264637, 18108364
2228	91227337 (4455, 4456)	Novel Protein sim. GBank gij806976 (U16800) - ribonucleoprotein [Xenopus laevis]	Contains protein domain (PF00076) - RNA recognition motif (a.k.a. RRM, RBD, or RNP domain)	dna_ma_bind	264091, 264092, 264094, 29331822, 29331825, 66714117, 264693, 263972, 264639, 83373044, 264563
2229	68060931 (4457, 4458)	Novel Protein sim. GBank gij3549155 (AC005625) - R27328_2 [Homo sapiens]		UNCLASSIFIED	

2230	95342915 (4459, 4460)	Novel Protein sim. GBank gij226154iprl14123504 - DNA polymerase [human adenovirus type 2]		UNCLASSIFIED	264488, 264768, 52644507, 264769, 21908765, 21906766, 21908787, 21908769, 22278995, 35695917, 22278996, 22278997, 22278998, 22278999, 265021, 264259, 52645129, 29331827, 284508, 284509, 284907, 18108370, 18108374, 35696423, 35696855, 265007, 264910, 264555, 35657402, 21906754, 18108387, 265010, 265018, 265019, 264760, 264266, 264567, 264583
2231	88060937 (4461, 4462)	Novel Protein sim. GBank gij3549154 (AC005625) - H27328.1 [Homo sapiens]		UNCLASSIFIED	
2232	87762581 (4463, 4464)	Novel Protein sim. GBank gij5261316ipjAAD41476:1(AFI13312) - (AFI133124) transcription factor. IIC83 [Homo sapiens]		transcription factor	18108394, 56182575, 22278995, 35696286, 22278997, 22278999, 284259, 29331827, 35696052, 264807, 56182435, 265008, 265007, 265008, 264910, 264758, 55812038, 284603, 265018, 265019, 18108351, 264682, 264784, 264683, 264369, 264288, 264688, 264687, 264689, 21906785, 21906766, 21906767, 21906769, 29148629, 35695917, 264690, 52844150, 284891, 33657023, 264693, 18108370, 18108374, 55811578, 35695855, 264639, 18108385, 264564, 284906, 33657402, 265018, 264288, 264888, 265020, 264835, 18108385
2233	87755282 (4465, 4466)	Novel Protein sim. GBank gij4249733ipjAAD13780] - (AFI09377) IidBP [Mus musculus]			56182435, 264369, 264688, 21906765, 285020, 264693, 264558, 56528468
2234	87771817 (4467, 4468)	Novel Protein sim. GBank gij1706559ipjP34352IEAS - DROME - ETHANOLAMINE KINASE (EASILY SHOCKED PROTEIN)		kinase	
2235	91012316 (4469, 4470)	Novel Protein sim. GBank gij4972734ipjAAD34762.1] - (AFI132174) unknown [Drosophila melanogaster]	Contains protein domain (PF00400) - WD domain, G-beta repeat	UNCLASSIFIED	22278997, 264563
2236	88003131 (4471, 4472)	Novel Protein sim. GBank gij1082675ipjIB53614 - p20 protein - human	Contains protein domain (PF00011) - eph Hsp20alpha crystallin family		264589, 264687, 264769, 265022, 264259, 60432049, 264691, 29331828, 60432289, 20281149, 284908, 264907, 264511, 265008, 285009, 264634, 284635, 284638, 284555, 284556, 264557, 284558, 60433359, 264595, 284559, 60433438, 60432113, 264761, 264762, 264763, 264764
2237	91012318 (4473, 4474)	Novel Protein sim. GBank gij4972734ipjAAD34762.1] - (AFI132174) unknown [Drosophila melanogaster]	Contains protein domain (PF00400) - WD domain, G-beta repeat	kinase	264488, 264769, 21908765, 21906768, 21908767, 21906768, 21906769, 22278995, 22278996, 22278997, 22278998, 265020, 265021, 22278999, 264259, 29331824, 29331826, 28331827, 35695763, 18108378, 35695855, 265007, 60432229, 33657402, 60433356, 60433438, 83373044, 18108385, 21906754, 18108387, 60432113, 22278900, 285019, 22279002, 264482, 18108351, 264286

2238	94898857 (4475, 4476)		Contains protein domain (PF00286) -		264509, 264907, 264629, 264634, 264564
2239	87786688 (4477, 4478)		Viral coat protein		
2240	84121471 (4479, 4480)	Novel Protein sim. GBank g12982311 (AF051240) - probable ubiquitin-conjugating enzyme E2 (Picea mariana)	Contains protein domain (PF00179) - Ubiquitin-conjugating enzyme	ubiquitin	29331825, 265009, 264389, 33657109, 18108370, 18108374, 264557, 264559, 264488, 65274572, 56182575, 35698286, 22278997, 22278999, 264259, 29331827, 35696052, 264508, 52644045, 56182435, 264571, 265007, 265008, 265009, 60433356, 60433438, 55612038, 21908754, 33657084, 55611386, 265018, 265019, 18108351, 264883, 264268, 264788, 264887, 264688, 264769, 21908765, 21908768, 21908769, 35695917, 265021, 265022, 60170815, 52844150, 33657023, 33657182, 33657349, 35695783, 18108370, 35698423, 35695855, 87168518, 22278000 264583, 264629
2241	80081851 (4481, 4482)	Novel Protein sim. GBank 812494312sp1705411E2BG_RAT - TRANSLATION INITIATION FACTOR EIF-28 GAMMA SUBUNIT (EIF-28 GDP-GTP EXCHANGE FACTOR)		UNCLASSIFIED synthase	22278995, 22278996, 22278997, 22278998, 264459, 29331822, 29331824, 29331828, 29331827, 29331828, 264509, 265007, 265008, 264598, 21908754, 265010, 265011, 265017, 265018, 265019, 264448, 264369, 264286, 52844229, 21908765, 21908766, 21908767, 21908768, 21908769, 265020, 265021, 33657109, 27488282, 27488284, 18108374, 35695855, 264834, 264637, 56182323, 63373044, 56526488, 87168518, 264564
2242	91228075 (4483, 4484)			UNCLASSIFIED	265008 264604
2243	78902028 (4485, 4486)	Novel Protein sim. GBank g12291143 (AF016417) - Similar to BZIP transcription factor (Caenorhabditis elegans)		UNCLASSIFIED	
2244	85723527 (4487, 4488)	Novel Protein sim. GBank g1470340 (U00043) - similar to beta-mannosyltransferase (Caenorhabditis elegans)		UNCLASSIFIED	
2245	85318545 (4489, 4490)		Contains protein domain (PF00534) - Glycosyl transferases group 1	UNCLASSIFIED	52645156, 22278995, 22278996, 22278997, 22278999, 29331822, 29331824, 29331827, 264907, 264512, 60433438, 264756, 21908754, 265011, 264603, 264764, 264687, 21908767, 21908768, 21908769, 55611957, 265022, 264691, 264628, 35698423, 264638, 18108387, 60432113, 22278000, 22279002, 264566

2246	94846710 (4481, 4482)	Novel Protein sim. GBank g14980696[dbj]AA76328.1 - (AB028068) activator of S phase Kinase [Homo sapiens]	Contains protein domain (PF00153) - Mitochondrial carrier proteins	transport	65274572, 22276995, 35696286, 22276996, 22276997, 22276999, 284259, 35696052, 264108, 264805, 264807, 265008, 265007, 265006, 60433438, 33108954, 67166559, 265018, 265019, 264288, 21906765, 21906767, 21906768, 21906769, 55611957, 35695917, 265020, 265022, 27466204, 16108370, 16106374, 65274791, 35696655, 60432113
2247	67862542 (4483, 4484)	Novel Protein sim. GBank g1464063[eml]CAA66337 - (X63413) U66 [Human herpesvirus 6]		UNCLASSIFIED	52645156, 52646365, 52645050, 35696052, 33656870, 52646317, 33657004, 265017, 21906766, 21906769, 35695917, 33657109, 52645129, 33657162, 27466261, 27466262, 33657349, 27466265, 16108367
2248	95412996 (4485, 4486)	Novel Protein sim. GBank g1476502[anp]_004123.1[ptABP - hyaluronan-binding protein 2]	Contains protein domain (PF00069) - Trypsin	cathepsin	264488, 264259, 264307, 26331630, 264909, 265007, 265009, 264595, 21906754, 65274444, 264803, 265019, 264782, 264448, 264288, 264688, 21906768, 55611957, 265021, 264891, 16108374, 264634, 264635, 264638, 264555, 264638, 264557, 264558, 264559, 16106363, 83373044, 16106385, 264486
2249	94685662 (4487, 4488)	Novel Protein sim. GBank g14038461 (AF107772) - TCS11 [Trypanosoma cruzi]	Contains protein domain (PF00515) - TPR Domain	eph	264766, 264626, 264638, 264637
2250	79627508 (4499, 4500)	Novel Protein sim. GBank g13736140[emb]CAA21241 - (AL031652) valyl-tRNA synthetase, mitochondrial precursor [Schizosaccharomyces pombe]		UNCLASSIFIED	264906, 16106374
2251	67365863 (4501, 4502)	Novel Protein sim. GBank g13218467[emb]CAA07090.1 - (AJ006529) putative phosphatase [Gallus gallus]		UNCLASSIFIED	264259, 35696052, 264508, 56162435, 265009, 264592, 264593, 264760, 264448, 264684, 264288, 264690, 264628, 55611576, 264555, 264556, 264557, 264558, 264559, 264566
2252	67735867 (4503, 4504)	Novel Protein sim. GBank g14928325[db]AAD33953.1[AF14531 - (AF145316) vacuolar proton pump della polypeptide [Homo sapiens]	Contains protein domain (PF01613) - ATP synthase subunit D	synthase	264092, 264094, 264259, 26331622, 66714117, 26331826, 264102, 264103, 264104, 264105, 264109, 264112, 264511, 265007, 60433358, 265010, 16108351, 21906767, 21906768, 264691, 263974, 263977, 264466, 264567
2253	81010703 (4505, 4506)			UNCLASSIFIED	65274572, 265019

2254	98320031 (4507, 4508)	Novel Protein sim. GBank g14502847[refNP_001271.1]CIRB - cold inducible RNA-binding protein	Contains protein domain (PF00076) - RNA recognition motif (a.k.a. RRM, RBD, or RNP domain)	dna_ma_bind	264569, 18108394, 18108396, 58182575, 58994075, 35696286, 22278999, 264084, 60432049, 264259, 29331822, 29331824, 29331825, 29331826, 60432289, 29331827, 29331828, 35696052, 264108, 264506, 264509, 264806, 264907, 29331630, 66712502, 264906, 264909, 264510, 265006, 264511, 265007, 265008, 265009, 60170631, 60432229, 60433358, 60433438, 264758, 65658542, 265010, 265011, 67168559, 265017, 265018, 265019, 264448, 264764, 264286, 264389, 264766, 264666, 264766, 264769, 21906785, 21906787, 55811937, 264691, 33657023, 264692, 18108362, 65274620, 263969, 264626, 18108370, 60431528, 263972, 264629, 18108372, 18108377, 18108378, 55811576, 35696423, 35696555, 264630, 264634, 264635, 264636, 264556, 263961, 264636, 58182323, 60170394, 264556, 18108381, 18108382, 63373044, 18108385, 67168518, 60432113, 22279002, 264482, 264564, 264565, 264466, 264567, 18108391
2255	91010546 (4508, 4510)	Novel Protein sim. GBank g15541665[emb CAB51072.1]- (AL098656) hypothetical protein [Homo sapiens]	Contains protein domain (PF00076) - RNA recognition motif (a.k.a. RRM, RBD, or RNP domain)	struct	65274572, 58182575, 22278997, 22278999, 264258, 29331822, 29331825, 29331826, 29331827, 29331828, 264506, 264905, 264908, 264907, 68712502, 264908, 56182435, 264510, 264511, 265008, 264593, 264595, 21906754, 33109954, 67168474, 265011, 265017, 265018, 264682, 264764, 264389, 264286, 264766, 264665, 264686, 264766, 21906765, 21906766, 21906769, 265020, 60170615, 52644150, 264690, 264692, 264693, 33657109, 33657349, 264632, 264636, 52644332, 56182323, 22279000, 22279002
2256	87020531 (4511, 4512)	Novel Protein sim. GBank g13327174[db BA31653]- (AB014560) KIAA0680 protein [Homo sapiens]		UNCLASSIFIED	264766, 264688, 18108374
2257	60088235 (4513, 4514)			UNCLASSIFIED	22278996, 22278999, 264661, 21906765, 21906766, 264567
2258	66090516 (4515, 4516)	Novel Protein sim. GBank g13025446 (AC004528) - R32184_2 [Homo sapiens]	Contains protein domain (PF00060) - Ligand-gated ion channel	misc_channel	264608, 264592, 264764

2259	95384155 (4517, 4518)	Novel Protein sim. GBank gij894140emb[CAB43378.1]- (AL050110) hypothetical protein [Homo sapiens]		UNCLASSIFIED	18108396, 65274572, 56182575, 22278997, 22278998, 264259, 29331822, 29331827, 264905, 66712502, 264908, 264909, 56182435, 285007, 285008, 80432229, 33857084, 87188559, 18108351, 264448, 264693, 264288, 264369, 56181582, 285021, 80170615, 264890, 33657109, 60431528, 18108374, 52644332, 56182323, 18108385, 22279000, 22279002, 264482
2260	88084119 (4519, 4520)	Novel Protein sim. GBank gij8080663 (ACD04614) - similar to f-eponidin proteins AB008088 (P1Dg2529225) [Homo sapiens]	Contains protein domain (PF00090) - Thrombospondin type 1 domain	oxidase	56182575, 265020, 264805, 264806, 264908, 33696423, 264511, 264635, 55812038, 264758, 265018, 265019, 264805, 264760, 264583
2281	88074157 (4521, 4522)	Novel Protein sim. GBank gij8334526emb[CAB418138] - (AL021308) predicted using FGENEH [Homo sapiens]		UNCLASSIFIED	56182575, 22278999, 29331822, 29331825, 80432289, 29331827, 35896052, 284508, 66712502, 52644045, 56182435, 265008, 265009, 265009, 60433358, 55812038, 265010, 265017, 285019, 264288, 264389, 21906765, 21906767, 55811957, 35895917, 52644150, 33657023, 33657109, 55811576, 65274791, 56182323
2283	87802495 (4523, 4526)	Novel Protein sim. GBank gij8341697 (ACD03872) - hypothetical protein [Arabidopsis thaliana]			22278994, 22278997, 264907, 264828, 52644150, 18108381, 264893, 18108374
2264	87756525 (4527, 4528)	Novel Protein sim. GBank gij1657601 (U66220) - unknown [Nemocrystis exedens]		UNCLASSIFIED	264688, 264486, 264768, 264769, 264691, 264508, 264905, 264509, 264908, 264907, 264908, 264909, 35695855, 264510, 264511, 264512, 265007, 285009, 264838, 264639, 264757, 264758, 18108385, 265011, 264760, 264584, 264585, 264784, 264586, 264486, 264768
2285	88918683 (4529, 4530)	Novel Protein sim. GBank gij477072jpr[AA8018] - mucin 7 precursor, salivary - human	Contains protein domain (PF00096) - Zinc finger, C2H2 type	UNCLASSIFIED	264689, 264810, 264784
2266	8773458 (4531, 4532)	Novel Protein sim. GBank gij3150479 (AF087212) - partial CDS [Caenorhabditis elegans]	Contains protein domain (PF01305) - Ribosomal protein L15 amino terminal region	ribosomal prot	22278995, 22278997, 22278999, 264259, 265006, 265007, 265009, 60433438, 21906754, 265010, 265011, 265017, 264448, 284883, 264288, 264889, 21908785, 21906768, 35695917, 265021, 18108374, 264638, 22279000, 22279002, 264588, 264487

2267	87395838 (4533, 4534)	Novel Protein sim. GBank g1380229[emb]CAA20697.11 - (AL031530) hypothetical protein [Schizosaccharomyces pombe]		UNCLASSIFIED	35686286, 264258, 28931824, 29331825, 35696052, 29331826, 264805, 264509, 264907, 264908, 264909, 264512, 285009, 264910, 264593, 33657402, 265010, 265018, 264782, 264448, 264288, 264369, 264786, 52644229, 35695917, 264691, 33657023, 18108382, 33657109, 35696423, 264634, 18108381, 87188518, 264568
2268	85683867 (4535, 4536)	Novel Protein sim. GBank g1726832[ep]P39189/ALU2_HUMAN - III ALU SUBFAMILY SB WARNING ENTRY III		cadherin	264488, 264259, 264508, 264595, 285010, 265017, 264766, 18108385, 264486
2269	86177877 (4537, 4538)	Novel Protein sim. GBank g1103416[pl]S17885 - TCD37 protein - fruit fly (Drosophila melanogaster)		UNCLASSIFIED	56182575, 60432048, 265007, 265009, 264591, 87188558, 264605, 18108351, 21906764, 265020, 264628, 60431528, 264838, 18108385, 18108387, 60432113, 264763
2270	80410327 (4539, 4540)			cyto450	264809, 56182435, 285008, 55812038, 55811957, 33857023, 264893, 33657108, 55810784, 55811578, 58182323
2271	81010392 (4541, 4542)			UNCLASSIFIED	264905, 264908
2272	84208220 (4543, 4544)	Novel Protein sim. GBank g1178370 (AC005058) - similar to calcium-independent phospholipase A2, similar to AC004392 (P1Dg3367519) (Homo sapiens)	Contains protein domain (PF00462) - Glutaredoxin	UNCLASSIFIED	52645158, 22278996, 22278999, 60432048, 264259, 29331822, 29331824, 29331825, 29331826, 29331827, 35696052, 264909, 265006, 264593, 60433438, 21908754, 285018, 264889, 21908765, 21906766, 21906767, 21906769, 265021, 265022, 60170815, 264691, 33857023, 264893, 33857109, 27488264, 18108376, 35696423, 35695855, 264630, 52644332, 264558, 58182323, 22279002
2274	91840217 (4547, 4548)	Novel Protein sim. GBank g11480112[emb]CAA679811 - (X98642) HP1-BP14 protein (Mus musculus)	Contains protein domain (PF00538) - linker histone H1 and H5 family	histone	52645156, 22278997, 22278999, 52645080, 29331824, 29331825, 29331826, 29331827, 29331828, 264905, 264908, 52644045, 264511, 265008, 265009, 60170831, 264591, 21908754, 33109854, 285011, 265018, 18108351, 264448, 264288, 264584, 264786, 21908785, 21906768, 21906767, 21906768, 52644150, 264683, 18108384, 35695783, 18108374, 35696423, 264634, 264557, 264836, 52644332, 63373044, 18108385, 56584868, 87188518, 22279002
2275	88082501 (4549, 4550)	Novel Protein sim. GBank g13165406 (AC004755) -	Contains protein domain (PF00122) -	transport	
2276	11297447 (4551, 4552)	Novel Protein sim. GBank g13165406 (AC004755) -	Contains protein domain (PF00122) -	transport	
2278	11297447 (4551, 4552)	Novel Protein sim. GBank g13165406 (AC004755) -	Contains protein domain (PF00122) -	transport	

2277	88084123 (4553, 4554)	Novel Protein sim. GBank gll2860078 (AC004142) - similar to murine leucine-rich repeat protein; possible role in neural development by protein-protein interactions; 93% similarity to D49802 (P10g138909) [Homo sapiens]	Contains protein domain (PF00560) - glycoprotein	22278999, 35696052, 285006, 285019, 264389, 285020, 285022, 55610784, 264404, 22279002
2278	94133078 (4555, 4556)	Novel Protein sim. GBank gll2616702 (AC002510) - unknown protein [Arabidopsis thaliana]	UNCLASSIFIED	18108394, 22278897, 28331628, 60433358, 6043438, 21906754, 285016, 33657023, 264639, 83373044, 264565
2279	60419375 (4557, 4558)	Novel Protein sim. GBank gll1197145pP13963EXTN_TORAC - EXTENSIN PRECURSOR (CELL WALL HYDROXYPROLINE-RICH GLYCOPROTEIN)	UNCLASSIFIED	264768, 264565
2280	94239723 (4559, 4560)	Novel Protein sim. GBank gll4240289 [d]BA74928, 11 - (AB020712) KIAA0905 protein [Homo sapiens]	Contains protein domain (PF00400) - transport WD domain, G-beta repeat	265008, 33109954, 265010, 285018, 265020, 264092, 264259, 29331622, 29331624, 29331828, 35696052, 264107, 264908, 264908, 52844045, 265008, 33657402, 60433358, 284756, 285011, 285018, 264881, 264663, 264664, 264688, 21906785, 21906787, 21906788, 21906789, 80170815, 284690, 52844150, 16108362, 264692, 18108368, 18108374, 263976, 264631, 18106381, 284559, 16108365, 56526486, 22279000, 264566, 264567
2282	67002828 (4563, 4564)	Novel Protein sim. GBank gll1537070 (U63840) - nucleoprotein p54 [Rattus norvegicus]	UNCLASSIFIED	264408, 264259, 29331822, 29331624, 29331627, 29331628, 29331630, 33657402, 60433436, 87168474, 285018, 18108351, 21906787, 21906789, 55611857, 33657023, 52845129, 33657109, 33657162, 27466282, 263972, 55611576, 87168518, 20281169
2283	95362386 (4565, 4566)	Novel Protein sim. GBank gll2495728 [p]Q92550 [Y281_HUMAN - HYPOTHETICAL PROTEIN KIAA0281 (HA8725)]	UNCLASSIFIED	80424179, 56182575, 22278894, 35696052, 22278987, 22278999, 29331622, 29331624, 56162181, 29331825, 29331827, 35696052, 28146498, 264905, 66712502, 264908, 265007, 265009, 60433228, 264593, 80431735, 60433356, 33109954, 33657084, 55611386, 87168474, 285010, 285011, 265018, 285019, 55811150, 264683, 264368, 264288, 264686, 21906785, 21906787, 21906788, 28146827, 21906789, 55611857, 265020, 265022, 33657162, 27468261, 18108370, 264628, 18108374, 55610784, 18108379, 55611576, 35696423, 35696555, 284630, 60431650, 263961, 18106382, 63373044, 18106385, 18106387, 60432113, 22279000, 264462, 264567

2284	85414955 (4587, 4588)	Novel Protein sim. GBank g 2498787 p C84311 PNAAD MOUSE - PROTEIN N-TERMINAL ASPARAGINE AMIDOHYDROLASE (PROTEIN NH2-TERMINAL ASPARAGINE DEAMIDASE) (NTN-AMIDASE) (PNAAD) (PROTEIN NH2-TERMINAL ASPARAGINE AMIDOHYDROLASE) (PNAAD)				60424179, 52644507, 18108394, 52846842, 22278994, 35698288, 22278998, 22278997, 22278999, 264259, 80432049, 29331822, 29331824, 29331825, 60432289, 29331826, 29331627, 35696052, 29331830, 52844045, 56182435, 33657402, 60433436, 33109954, 21906754, 85658542, 87168559, 265016, 265019, 55811150, 264682, 264369, 21906765, 21908768, 21808767, 21806786, 21806789, 55811957, 35695917, 265020, 265021, 60170815, 33857023, 33857162, 27486262, 27486264, 27486285, 18108376, 55810764, 35696423, 35695855, 60431850, 87168518, 80432113, 264462, 264594
2285	87781484 (4569, 4570)	Novel Protein sim. GBank g 3342234 (U83909) - nuclear antigen EBNA-1 [Cercophthodina herpesvirus 1s]			collagen	35696052, 264505, 264807, 264308, 264809, 264512, 265009, 264810, 264595, 264760, 18106351, 264882, 264783, 264685, 264766, 264686, 264768, 264683, 264629, 35695855, 264631, 264634
2286	87737625 (4571, 4572)	Novel Protein sim. GBank g 3873414 (U00043) - similar to D. melanogaster trithorax protein [Caenorhabditis elegans]			kinase	35696286, 58182435, 60170831, 264591, 60432229, 264592, 264593, 264594, 264595, 55812038, 264596, 67168474, 35695917, 264692, 55811576, 264555, 264557
2287	82868896 (4573, 4574)	Novel Protein sim. GBank g 230905 p S42731 - collagen alpha 1 chain - sea urchin (Hemicentrotus pulcherrimus) (fragment)			UNCLASSIFIED	264682
2288	84133063 (4575, 4576)	Novel Protein sim. GBank g 1728632 p P39189 ALU2_HUMAN - III ALU SUBFAMILY SB WARNING ENTRY IIII			Im7	22278985, 56994075, 22278997, 22278999, 264259, 60432289, 264508, 264512, 265008, 33657402, 265017, 265018, 265019, 18108351, 264448, 21908765, 21908768, 21908787, 21908788, 33657023, 264557, 22279000, 22279002
2289	88084133 (4577, 4578)	Novel Protein sim. GBank g 2867497 (AC004144) - R34001_1 [Homo sapiens]			kinase	
2290	88084137 (4579, 4580)	Novel Protein sim. GBank g 2887497 (AC004144) - R34001_1 [Homo sapiens]			UNCLASSIFIED	264683
2291	84295281 (4581, 4582)	Novel Protein sim. GBank g 3253120 (AC005175) - R31449_3 [Homo sapiens]			struct	18108394, 264807, 265006, 265009, 33109954, 52646317, 265010, 16108331, 264881, 264688, 264692, 18108370, 16108374, 16108385

2292	94328834 (4583, 4584)	Novel Protein sim. GBank gi14803672[emb]CAB2843.11 - (Al133789) nuclear transport receptor [Homo sapiens]		UNCLASSIFIED	56182575, 35686286, 56894075, 29331824, 29331825, 35898052, 56182435, 60433436, 55612036, 33109954, 67168474, 67168559, 265018, 16106351, 264763, 264446, 264369, 264286, 56181562, 264769, 21906785, 21906766, 21906767, 21906769, 265021, 265022, 33857023, 264693, 65274620, 33657109, 27486264, 284629, 55610764, 55611576, 35695855, 56182323, 56528466, 67168518, 22279000, 264567
2293	67759213 (4585, 4586)	Novel Protein sim. GBank gi13252981 (AF068921) - Ras-binding protein SURF-6 [Mus musculus]	Contains protein domain (PF00560) - Ras-Leucine Rich Repeat	strucd	264488, 18108397, 35696286, 264092, 264259, 29331822, 29331826, 264906, 264906, 264511, 264512, 265009, 264910, 16108351, 264764, 264389, 264286, 264655, 264768, 265020, 265022, 264534, 35696423, 264831, 264637, 18108381, 56182323, 264639, 18108385, 264404, 264563, 264565
2294	86893580 (4587, 4588)	Novel Protein sim. GBank gi12082680 (U88964) - HEM45 [Homo sapiens]	Contains protein domain (PF00929) - nuclease Exonuclease	nuclease	22276997, 22276998, 22278999, 264259, 29331822, 60432289, 29331826, 35696052, 265018, 264684, 264286, 264688, 21908785, 21906766, 21906767, 21908768, 21908769, 265020, 265021, 264682, 33657109, 16106376, 35696423, 35695855, 284634, 22279000, 22279002, 284563, 284486
2295	95312200 (4589, 4590)			UNCLASSIFIED	22278996, 60432288, 264682, 264683, 264889, 18108374
2296	80030781 (4591, 4592)				263974, 263978
2297	94321251 (4593, 4594)	Novel Protein sim. GBank gi15689501[id]BAAB3034.11 - (AB029005) KIAA1082 protein [Homo sapiens]		transcriptador	264468, 65274572, 56182575, 22278997, 22278998, 264259, 29331822, 29331824, 29331826, 29331828, 35898052, 264907, 264908, 52644045, 56182435, 264112, 265006, 265007, 264910, 265009, 60433356, 33857402, 264595, 55612036, 21908754, 265011, 265018, 265019, 264446, 264764, 264286, 264768, 21908765, 21908767, 21906766, 21908768, 55611857, 265020, 265021, 264690, 264691, 33857023, 264692, 264893, 33857109, 55611576, 56182323, 60170394, 63373044, 16108365, 56528466, 264564, 264486

2286	85312207 (4595, 4596)	Novel Protein sim. GBank gij3675051[emb]CA802849) - (261050) predicted using GeneIndex; similar to collagen; cDNA EST EMBL.D65584 comes from this gene; cDNA EST EMBL.D68046 comes from this gene; cDNA EST yk368b12.3 comes from this gene; cDNA EST yk368b12.5 comes from this gene ...		collagen	60424178, 56161686, 22276995, 35686266, 22276996, 22276998, 22276999, 264490, 264259, 29331622, 29331824, 66714117, 60424269, 35686052, 29331826, 68712502, 56162435, 264510, 265006, 60433436, 21806734, 33109934, 55811386, 265010, 265016, 55811150, 264782, 16106331, 264682, 264683, 264268, 264684, 264688, 264686, 56181562, 264689, 21806766, 21806767, 29148629, 55811957, 29146764, 35695917, 265020, 16108362, 33657023, 16108364, 33657109, 60431602, 16106370, 60431526, 16108374, 55810764, 35696423, 35695855, 264630, 264634, 60431650, 16106380, 56182323, 264556, 63373044, 16106385, 60432113, 22279000, 264462, 264567, 264466
2289	60193720 (4597, 4598)	Novel Protein sim. GBank gij2443866 (AC002294) -		UNCLASSIFIED	264389
2300	84124346 (4598, 4600)	Unknown protein [Arabidopsis thaliana]		UNCLASSIFIED	264466, 22276998, 22276999, 264259, 29331624, 66714117, 35686052, 264509, 264805, 264806, 264807, 264808, 264809, 265008, 264810, 265009, 264756, 265010, 67168559, 264600, 265016, 264760, 264762, 16106351, 264764, 264766, 264768, 264769, 21806768, 21806767, 35695917, 265021, 264691, 33657023, 35695763, 16106370, 16108374, 35696423, 35695855, 264631, 264636, 264638, 16106385, 22278002, 264563
2301	91235725 (4601, 4602)	Novel Protein sim. GBank gij2143637[pil]64505 - calcium-dependent actin-binding protein - rat		struct	264908, 264758, 265017, 21806765, 83373044, 264563
2302	86084141 (4603, 4604)	Novel Protein sim. GBank gij2887497 (AC004144) -		UNCLASSIFIED	52644045, 265018, 264288, 33657023, 16106370, 16106385
2303	94141439 (4605, 4606)	Novel Protein sim. GBank gij4884194[emb]CA843220 - 1 - (ALM9848) hypothetical protein [Homo sapiens]		struct	264259, 80432048, 264807, 264809, 264810, 60432229, 33657402, 265011, 265016, 264762, 264446, 264769, 264637, 264638, 83373044, 264466
2304	94640434 (4607, 4608)	Novel Protein sim. GBank gij2484162[sp]Q10005[RYR1, CAEL - HYPOTHETICAL 39.9 KD PROTEIN T15H9.1 IN CHROMOSOME II PRECURSOR		UNCLASSIFIED	264259, 29331624, 21806767, 33657162, 33657349
2305	90835911 (4609, 4610)	Novel Protein sim. GBank gij4872686[sp]AA034738 - 1 - (AF132150) unknown [Drosophila melanogaster]			65274572, 22276998, 264908, 265006, 21906768, 264681, 264486

2308	95334840 (4611, 4612)	Novel Protein sim. GBank g1482855ipjAAD34043.1/AF15180 - (AF151808) CGI-48 protein (Homo sapiens)	Contains protein domain (PF00400) - WD domain, G-beta repeat	Kinase receptor	264488, 22278995, 22278996, 35898266, 22278997, 22278998, 22278999, 264259, 29331622, 29331624, 29331825, 29331828, 29331827, 29331828, 35896052, 264508, 264905, 264907, 29331830, 264908, 264909, 264510, 285006, 264511, 265007, 264512, 285006, 285009, 33657402, 21906754, 85658542, 265010, 265011, 264801, 265017, 265016, 264604, 265019, 18108351, 264446, 264288, 264768, 264789, 21908765, 21908768, 21906767, 21906768, 29148628, 29148784, 35895917, 265020, 265021, 265022, 33657023, 264892, 18108370, 16108374, 18108378, 35898423, 35895855, 264630, 264634, 264635, 264636, 264637, 264638, 264639, 18108382, 18108385, 18108387, 264583, 284566, 264486, 264828
2307	79415263 (4613, 4614)	Novel Protein sim. GBank g14758732ipfNP_004522.1/mOCS - molybdenum cofactor synthase 2		UNCLASSIFIED synthase	35896286, 264259, 29331822, 29331824, 264112, 264512, 264757, 21908754, 264288, 264690, 27486264, 264631, 284634, 264404
2306	87608409 (4615, 4616)	Novel Protein sim. GBank g13678059emb CAB17070 - (Z99942) cDNA EST EMBL-D73444 comes from this gene; cDNA EST EMBL-D70905 comes from this gene; cDNA EST EMBL-D72208 comes from this gene; cDNA EST EMBL-D75030 comes from this gene; cDNA EST EMBL-D72944 comes from this gene; cDN...		glycoprotein	18106387, 22278998, 22278997, 22278996, 22278999, 80432049, 29331822, 29331826, 60432289, 66712502, 60432229, 60433356, 80433438, 65274444, 265010, 264600, 264881, 264448, 264683, 264288, 21908766, 21908768, 285020, 264891, 264692, 264693, 65274620, 85274791
2310	78801866 (4619, 4620)	Novel Protein sim. GBank g12137337ip l 48281 - gene		UNCLASSIFIED	264508
2311	87721189 (4621, 4622)	mCBP protein - mouse	Contains protein domain (PF00013) - KH domain	transcript factor	18108397, 56182575, 22278996, 56894075, 264259, 29331824, 29331827, 264508, 264907, 56182435, 264510, 264511, 285008, 264512, 265007, 265008, 285009, 60433438, 33109954, 265010, 265011, 264603, 265017, 16106351, 264782, 264863, 264288, 264369, 284666, 33657023, 20281149, 20281069, 264628, 263972, 55811576, 35898423, 20261071, 264632, 264638, 18108385, 18108387, 67168518, 22279000, 264583, 264486

2312	87549881 (4823, 4624)	Novel Protein sim. GBank gij2911264 (AC002550) - Unknown gene product [Homo sapiens]			5618257, 56994075, 35686288, 22278996, 22278997, 22278999, 264259, 20331822, 29331824, 29331825, 29331826, 28331827, 52644045, 60432229, 60433358, 55812038, 33109954, 21906754, 87168474, 265018, 18108351, 264288, 52644229, 21908765, 21906768, 21906767, 21906768, 21906769, 265022, 80170815, 33857023, 27488261, 27486264, 35696423, 35695855, 18108385, 22279000, 22279002
2313	80042533 (4825, 4826)	Novel Protein sim. GBank gij3043626dhljBA4254771 - (AB011123) KIAA0551 protein [Homo sapiens]			283981
2314	94313401 (4627, 4828)	Novel Protein sim. GBank gij5598714lambjCAB51401.11 - (AL035398) dJ796117.2 (CCF-51) [Homo sapiens]		UNCLASSIFIED	52644507, 52646365, 52640842, 22278996, 22278997, 22278999, 264259, 52645080, 29331822, 29331824, 60432229, 29331826, 29331827, 29331828, 35696052, 33658970, 52644045, 285008, 264593, 80433358, 60433438, 264758, 33109954, 285010, 285017, 285018, 265019, 264288, 264369, 21908765, 21906768, 21906768, 35695917, 52644150, 33857023, 33857109, 52845129, 33657349, 35695763, 18108374, 35696423, 35695855, 52644332, 22279000, 22278002, 264563, 264567
2315	80430119 (4629, 4630)			UNCLASSIFIED	264905, 264806, 264767, 264768, 264693, 55811578, 264635, 56182323, 18108385
2316	94312181 (4631, 4632)	Novel Protein sim. GBank gij5531827jbljAAD44486.11 - (AF078656) p47 [Homo sapiens]	Contains protein domain (PF00769) - UBX domain	glycoprotein	52644507, 52645156, 52648385, 22278994, 22278995, 22278996, 56994075, 35696286, 22278997, 22278998, 22278999, 264259, 29331822, 52645080, 29331824, 29331825, 29331826, 29331827, 35696052, 29331828, 264908, 264907, 29331830, 52844045, 58182435, 264511, 265007, 265008, 265009, 80170831, 80433438, 21908754, 52646317, 33109954, 33857084, 52844298, 87168474, 265010, 87168559, 265017, 285018, 265019, 264681, 264763, 264448, 264683, 264389, 52644229, 21906764, 264889, 21906765, 21906768, 21906767, 21908768, 21906769, 35695917, 285020, 285021, 52844150, 33657023, 18108362, 52645129, 33657182, 33657349, 35695763, 18108370, 18108378, 35696423, 35695855, 264831, 264558, 52644332, 83373044, 18108385, 18108387, 87188518, 80432113, 22278000, 264568, 264567

2317	87020571 (4633, 4634)			UNCLASSIFIED	22278998, 60432048, 264810, 60432228, 264686, 264687, 264688, 264689, 264558, 18108385
2318	79859678 (4635, 4636)			UNCLASSIFIED	264488, 264569, 18108396, 52846385, 22278994, 22278995, 22278996, 56994075, 35686286, 22278997, 22278998, 264239, 52845080, 28331825, 28331826, 28331827, 29331828, 29331830, 56182435, 60170831, 60432229, 80431735, 33657402, 21806754, 52844298, 87188474, 265011, 87168559, 265017, 265018, 265019, 18108351, 284448, 18108354, 264288, 284369, 52844229, 21906764, 21906765, 21906766, 21906767, 21906768, 21906769, 265021, 265022, 52844150, 33657023, 52845129, 33657109, 27486284, 33657349, 35685763, 18108370, 18108376, 18108379, 35686423, 284558, 83373044, 18108385, 56528486, 87168518, 264584, 264585, 264586
2320	91622426 (4639, 4640)	Novel Protein sim. GBank glij28637j p39194 ALU_HUMAN - III ALU SUBFAMILY SQ WARNING ENTRY III		Kinase	22278994, 60432049, 60432289, 28331827, 264511, 265008, 52846317, 265017, 265018, 21906765, 18108372, 18108387, 22279002
2321	94320377 (4641, 4642)	Novel Protein sim. GBank glij3873837 emb CAB02700 - (261028) Similarity to S.pombe hypothetical protein C1D4.09C (SW.Q10154). cDNA EST EMBL.T00543 comes from this gene. cDNA EST EMBL.T01062 comes from this gene. cDNA EST EMBL.T01321 comes from this gene. cDNA EST EMBL.T02288 com...		UNCLASSIFIED	284408, 264687, 18108394, 264688, 21906765, 18108387, 18108388, 21906767, 21908768, 65274791, 22278995, 35695855, 22278998, 265021, 265022, 264510, 265008, 264511, 264512, 265008, 60170815, 284555, 284636, 284556, 18108361, 264239, 60432229, 33657023, 284557, 284558, 264893, 60433356, 264559, 60433438, 29331824, 18108365, 18108348, 18108384, 29331825, 18108385, 33109954, 28331827, 56526486, 28146489, 265011, 60432113, 265017, 265018, 284508, 284563, 284482, 284509, 18108351, 284448, 284907, 264682, 18108370, 264683, 264908, 264288, 264909, 18108354, 264488, 264567
2322	87803165 (4643, 4644)	Novel Protein sim. GBank glij587857 emb CAB51685.1 - (AL109830) BACRTA.y [Drosophila melanogaster]	Contains protein domain (PF00106) - short chain dehydrogenase	- dehydrogenase	22278996, 264907, 264511, 264757, 18108351, 284768, 264838

2323	94840445 (4645, 4646)	Novel Protein sim. GBank glij284162lepiQ100051RYV1. CAEL - HYPOTHETICAL 39.9 KD PROTEIN T15H8.1 IN CHROMOSOME II PRECURSOR	Contains protein domain (PF00226) - eph DnaI domain		22278984, 22278995, 22278997, 60432049, 264259, 29331822, 33656970, 264509, 58182435, 284511, 265008, 60433358, 60433438, 55812038, 33109954, 21906754, 65858542, 87168474, 285011, 87168559, 265017, 285019, 264760, 264681, 16108351, 284368, 264288, 18108355, 284887, 284688, 21906765, 21906767, 21906768, 55811957, 35685917, 265021, 33857023, 16108362, 27486262, 55811576, 264631, 264555, 63373044, 87188518, 60432113, 22279002 264592, 284583, 265020
2324	86633607 (4647, 4648)	Novel Protein sim. GBank glij5419865lemb/CAB46377.1 - (AL096732) hypothetical protein [Homo sapiens]		ATPase, associated	265020
2325	88185074 (4649, 4650)	Novel Protein sim. GBank glij231885sp/P28801(CP4C, BLAD) - CYTOCHROME P450 4C1 (CYP1VC1)	Contains protein domain (PF00067) - cyto450 Cytochrome P450		265006, 264758, 35685855, 58182323
2326	84330862 (4651, 4652)	Novel Protein sim. GBank glij231885sp/P28801(CP4C, BLAD) - CYTOCHROME P450 4C1 (CYP1VC1)		UNCLASSIFIED	
2327	88081648 (4653, 4654)	Novel Protein sim. GBank glij4240227[idb]BA474892.1 - (AB020678) KIAA0869 protein [Homo sapiens]		UNCLASSIFIED	264259, 284508, 264905, 264806, 264807, 264808, 265007, 264512, 264910, 264758, 285010, 284768, 264768, 264769, 33857023, 284893, 284628, 264631, 264634, 264638, 284538, 264488
2328	83388428 (4655, 4656)	Novel Protein sim. GBank glij245105 (U46463) - glutamine repeat protein-1 [Mus musculus]		UNCLASSIFIED	60433438, 264595, 285017, 284766, 284692, 284828, 264635, 264638, 264638, 56182323, 60432113, 264566
2329	87604476 (4657, 4658)	Novel Protein sim. GBank glij1169343sp/P42209(DIPE, MOUSE - DIPE6 PROTEIN	Contains protein domain (PF00735) - UNCLASSIFIED Cell division protein		265017, 264685, 60432113, 284088 265009
2330	87335386 (4659, 4660)	Novel Protein sim. GBank glij5679138b/AD4874.1(AF16093) - (AF160934)		UNCLASSIFIED	
2331	86990463 (4661, 4662)	Novel Protein sim. GBank glij5679138b/AD4874.1(AF16093) - (AF160934)		transport	
2332	87784162 (4663, 4664)	Novel Protein sim. GBank glij2104452lemb/CAB08779) - (Z85397) unknown [Schizosaccharomyces pombe]		ATPase, associated	35686286, 22278998, 29331824, 60424269, 265006, 265008, 285018, 264448, 284764, 21906765, 35685917, 35685855, 264636, 22279000, 284566
2333	86206958 (4665, 4666)	Novel Protein sim. GBank glij387995lemb/CAA92691.1 - (Z68318) cDNA EST CEWSDE2F comes from this gene: cDNA EST EMBL:CO7930 comes from this gene: cDNA EST EMBL:CO9483 comes from this gene: cDNA EST y44156a.3 comes from this gene: cDNA EST y44156a.5 comes from this gene: cDNA EST ...	UNCLASSIFIED		56182575, 28331825, 21908768, 264638, 284107, 33857402, 87168559, 264683, 35685917, 285021, 33857023, 203976
2334	84319788 (4667, 4668)	Novel Protein sim. GBank glij4986270b/AA652261.21 - (U97002) similar to acyl-CoA dehydrogenases and epoxide hydrolases; Pfam domain PF00441 (Acyl-CoA dh), Score=57.4, E-value=1.7e-18, N=2; contains similarity to Pfam domain PF00702 (hydrolase), Score=57.4, E- value=1e-13, N=1 [C...]	Contains protein domain (PF00441) - dehydrogenase Acyl-CoA dehydrogenase		56182575, 28331825, 21908768, 264638, 83373044

2335	80046103 (4689, 4670)	Novel Protein sim. GBank gl 3283350 (AF062378) - calmodulin-binding protein SHA1 [Mus musculus]	Contains protein domain (PF00612) - struct	16106351, 21906789, 264555
2336	95196121 (4671, 4672)	Novel Protein sim. GBank gl 1928056 (emb CAA72805.1) (Y12090) putative 3,4-dihydroxy-2-butenone kinase [Lycopersicon esculentum]	kinase	264907, 35695917, 16106379
2337	95345610 (4673, 4674)	Novel Protein sim. GBank gl 4495063 (emb CA839181.1) (Z85986) dJ106K11.3 (similar to yeast suppressor protein SRP40) [Homo sapiens]	UNCLASSIFIED	35696286, 22276999, 56182181, 29331825, 60424269, 56182435, 33657402, 55812038, 55811388, 265017, 265018, 265019, 21906786, 35695917, 264691, 33657023, 33657108, 263972, 35696423, 35695855, 60432113
2338	87634045 (4675, 4676)	Novel Protein sim. GBank gl 2224689 (db BAA20629) - (AB002372) KIAA0374 [Homo sapiens]	Contains protein domain (PF00323) - Mammalian defensin	UNCLASSIFIED
2339	85663318 (4677, 4678)	Novel Protein sim. GBank gl 3673550 (emb CAA22127) - (AL033534) serine-rich protein [Schizosaccharomyces pombe]	UNCLASSIFIED	35696286, 264592, 264389, 264691, 264556
2340	90937718 (4678, 4680)			65274572, 22278894, 35696286, 22276997, 22276999, 264258, 29331822, 60432289, 29331826, 29331830, 265009, 33657402, 33109954, 265017, 265018, 264786, 264685, 21906789, 35695917, 264691, 264692, 35696423, 87168516, 22279000
2341	67775281 (4681, 4682)			264259, 264808, 264809, 264862, 22279000
2342	95334868 (4683, 4684)	Novel Protein sim. GBank gl 3673553 (emb CA802797) - (Z81042) similar to Yeast hypothetical protein YEV6 like: cDNA EST YK206h5.3 comes from this gene: cDNA EST YK206h5.5 comes from this gene: cDNA EST YK303h1.3 comes from this gene: cDNA EST YK303h1.5 comes from this gene: cDNA ...	Contains protein domain (PF00400) - VWD domain, G-beta repeat	kinase
2343	87775448 (4685, 4686)	Novel Protein sim. GBank gl 4929741 (db AAD34131.1) (AF15189) - (AF151894) CGI-136 protein [Homo sapiens]	UNCLASSIFIED	264907, 264512, 265011, 264683
2344	70953198 (4687, 4688)		UNCLASSIFIED	264758
2345	94318798 (4689, 4690)	Novel Protein sim. GBank gl 2506307 (db P13944) (CAIC - CHICK - COLLAGEN ALPHA 1(XII) CHAIN PRECURSOR (FIBROCHIMERIN)	Contains protein domain (PF00082) - collagen von Willebrand factor type A domain	264488, 264259, 66712502, 264759, 63373044, 264566

2346	94131820 (4691, 4692)	Novel Protein sim. GBank gll1255411 (U53153) - one short region of weak similarity to <i>S. cerevisiae</i> protease A inhibitor 3 (SP:P01094) and another short region of weak similarity to <i>S. cerevisiae</i> glucose repression mediator protein (SP:P14922) [<i>Caenorhabditis elegans</i>]	Contains protein domain (PF00515) - TPR Domain	proteaseinhib	35696286, 22276998, 264259, 35696052, 29331826, 33657402, 60433336, 33109984, 87166359, 264803, 265019, 18108351, 284661, 264685, 21908766, 285021, 33657109, 55811576, 35895655, 284637, 52644332, 264557, 83373044, 22279000, 22279002
2347	8530367 (4693, 4694)				
2348	95188133 (4695, 4696)	Novel Protein sim. GBank gll18280561emb1(CAA72805) - (Y12080) putative 3,4-dihydroxy-2-butanone kinase [Lycopodium obscurum]		kinase	18108394, 35696286, 264259, 35696052, 264508, 264509, 264905, 264906, 264907, 264908, 264909, 264510, 264511, 265006, 265007, 264512, 265008, 265009, 264910, 264591, 264592, 264593, 264594, 264757, 264595, 264596, 264758, 265011, 264601, 264762, 18108351, 264764, 264288, 264766, 264766, 264689, 35695917, 264893, 264628, 18108370, 264829, 18108374, 35696423, 264631, 264635, 264638, 264637, 264836, 264639, 63373044, 18108385, 264567, 264486
2349	8776502 (4697, 4698)	Novel Protein sim. GBank gll1061emb1(CAB43254.1) - (AL050062) hypothetical protein [Homo sapiens]			35696052, 29146489, 264909, 284389
2350	88260594 (4699, 4700)				22276998, 22276999, 264259, 29331822, 29331824, 29331825, 29331827, 29331828, 33109984, 21906754, 265010, 87166558, 265018, 265019, 284781, 264681, 264286, 18106357, 21906766, 21906767, 264691, 264692, 35695855, 87166518, 22279000, 22279002, 264482
2351	66968042 (4701, 4702)	Novel Protein sim. GBank gll728832isplP39189ALU2_HUMAN - IIII ALU SUBFAMILY S8 WARNING ENTRY IIII	kinase		56182575, 264909, 265006, 264558
2352	87337198 (4703, 4704)	Novel Protein sim. GBank gll731637isplP38760YHH5_YEAST - HYPOTHETICAL 75.9 KD PROTEIN IN SPO13-ARG4 INTERGENIC REGION	Contains protein domain (PF00076) - RNA recognition motif, (a.k.a. RRM, RBD, or RNP domain)	UNCLASSIFIED	264259, 264448
2353	91638784 (4705, 4706)	Novel Protein sim. GBank gll1346955isplP48609R827_DROME - HETEROGENEOUS NUCLEAR RIBONUCLEOPROTEIN 27C (HNRNP 48) (HRP48.1)	Contains protein domain (PF00076) - RNA recognition motif, (a.k.a. RRM, RBD, or RNP domain)	dna_ma_bind	29331826, 55812038, 265018, 264692, 264636
2354	87337198 (4707, 4708)	Novel Protein sim. GBank gll731637isplP38760YHH5_YEAST - HYPOTHETICAL 75.9 KD PROTEIN IN SPO13-ARG4 INTERGENIC REGION	Contains protein domain (PF00076) - RNA recognition motif, (a.k.a. RRM, RBD, or RNP domain)	UNCLASSIFIED	29331824, 264906, 265006, 265008

2355	91838786 (4709, 4710)	Novel Protein sim. GBank gl1433553jenn[CABA3861.1] - (AL078485) hnRNP-like protein [Arabidopsis thaliana]	Contains protein domain (PF00076) - dms_ma_bind RNA recognition motif (a.k.a. RRM, RBD, or RNP domain)		56994075, 22278996, 35696286, 22278999, 264259, 29331825, 29331826, 29331828, 29146498, 284905, 264906, 265006, 264756, 67168474, 285010, 265017, 264687, 21908785, 21908767, 21906789, 284691, 284692, 283967, 18108370, 67168518, 22279000
2356	95327688 (4711, 4712)	Novel Protein sim. GBank gl15138920[gbpAAD0377.1] - (AF092135) PTD014 [Homo sapiens]			52644507, 22278995, 35696286, 22278996, 22278997, 22278998, 22278999, 284259, 29331824, 68714117, 29331625, 60432289, 35696052, 29331828, 284906, 66712502, 264512, 285007, 265008, 60170831, 60432229, 60433356, 60433438, 284758, 52648317, 33109954, 21906754, 55811308, 87168474, 285017, 265018, 264605, 265019, 264681, 264682, 264448, 284369, 264288, 284686, 264768, 21906765, 21906766, 21906767, 21906768, 21906769, 285021, 60170815, 33657109, 27488284, 35695763, 55810764, 18108378, 35696423, 55611576, 35895855, 60170394, 56182323, 83373044, 18108365, 56526468, 264404, 80432113, 22279000, 264482, 264563, 264588, 264486, 264567
2357	87775458 (4713, 4714)	Novel Protein sim. GBank gl162829741[gbpAAD04131.1] (AF15189 - (AF151894) CGH-136 protein [Homo sapiens])		UNCLASSIFIED	284488, 264769, 16108394, 264259, 29331822, 16108370, 18108374, 264510, 265017, 264482, 264563, 264782, 264565, 264566, 264369, 18108354
2358	87777076 (4715, 4716)	Novel Protein sim. GBank gl14216005 (AC006135) - putative vicilin storage protein (globulin-like) [Arabidopsis thaliana]		UNCLASSIFIED	22278997, 22278998, 264509, 264805, 264592, 18108351, 264681, 264882, 264769, 32833988, 18108374, 264556, 18108365, 264482
2359	87755859 (4717, 4718)	Novel Protein sim. GBank gl1086830 (U41264) - coded for by C. elegans cDNA yk2016.5; coded for by C. elegans cDNA yk4491.5; coded for by C. elegans cDNA yk3686.5; coded for by C. elegans cDNA yk2016.5; coded for by C. elegans cDNA yk18912....		UNCLASSIFIED	35696286, 22278998, 264805, 264511, 265007, 265008, 60433438, 264286, 264686, 21906769, 265020, 264692, 35695855, 264558, 56526486, 264563
2360	80046125 (4719, 4720)	Novel Protein sim. GBank gl13861545[emb]CAA93779] - (Z69904) cDNA EST yk428d5.3 comes from this gene; cDNA EST yk428d5.5 comes from this gene [Caenorhabditis elegans]		UNCLASSIFIED	22278997, 29331826, 263961, 22279000
2361	94232181 (4721, 4722)	Novel Protein sim. GBank gl1746467 (U23514) - No definition line found [Caenorhabditis elegans]			22278995, 22278999, 264512, 285009, 264757, 21906765, 65274620, 18108370, 80431528, 18108374, 284635, 60170394, 264482

2382	81721193 (4723, 4724)	Novel Protein sim. GBank gll171083sp187081MYSB.ACACA - MYOSIN HEAVY CHAIN IB (MYOSIN HEAVY CHAIN II)		UNCLASSIFIED	22276898, 264258, 29331822, 29331824, 60432289, 264509, 264512, 60432229, 60433358, 264448, 264682, 264683, 264369, 21908765, 21906768, 21906769, 60432113, 22279000, 22279002
2383	85006835 (4725, 4728)	Novel Protein sim. GBank gll54055iemjCAAS8337 - (XG3413) U88 [Human herpesvirus 8]		UNCLASSIFIED	264907, 264629, 264635
2384	84827104 (4727, 4728)	Novel Protein sim. GBank gll539830jblAD45868.1AF14601 - (AF146016) hydroxypyruvate reductase [Homo sapiens]	Contains protein domain (PF00388) - reductase D-isomer specific 2-hydroxyacid dehydrogenases		264488, 18108394, 264887, 18108398, 22276996, 56994075, 35696286, 22276987, 22276998, 264259, 66714117, 29331825, 35898052, 264509, 264905, 264906, 264907, 264908, 66712502, 264909, 264511, 265008, 264512, 265007, 265008, 33657402, 264758, 21906754, 87168474, 265010, 87168559, 264603, 265017, 265018, 265019, 264760, 264762, 18108351, 264448, 264784, 264683, 264684, 264288, 18108355, 264766, 18108358, 264889, 18108359, 21906785, 21906786, 21906787, 35895917, 265020, 265021, 265022, 80170815, 52644150, 264691, 33657023, 264692, 18108354, 33657109, 18108368, 18108370, 18108374, 35698423, 35698855, 264635, 264556, 264557, 264639, 60170394, 83373044, 18108383, 18108384, 18108385, 18108386, 56526488, 264482, 264584, 264488
2385	94140746 (4729, 4730)	Novel Protein sim. GBank gll1840045 (U49082) - transporter protein [Homo sapiens]		transport	22276996, 22276998, 22276999, 264907, 264908, 264910, 33657402, 264758, 264600, 264786, 264687, 264889, 21906785, 21906787, 21906788, 21906789, 285021, 33657023, 33657109, 83373044, 264588
2386	84312386 (4731, 4732)			UNCLASSIFIED	52644507, 284259, 29331824, 29331825, 29331828, 29331827, 29331828, 284807, 29331830, 264909, 264511, 265008, 33857402, 264595, 52646317, 265017, 265018, 265019, 284605, 264685, 264766, 264689, 21906768, 21906769, 35695917, 265020, 265021, 265022, 52644150, 35695855, 52644332, 18108385, 18108387, 264564, 264588
2387	84140910 (4733, 4734)	Novel Protein sim. GBank gll1065457 (U40410) - C54G7.4 gene product [Caenorhabditis elegans]	Contains protein domain (PF00400) - WD domain, G-beta repeat	UNCLASSIFIED	35696286, 21906766, 55810764, 65274791, 264567
2388	84322190 (4735, 4736)				284628

2369	94314334 (4737, 4738)	Novel Protein sim. GBank gll580901[db]BAA62156.11 - (AB026343) a-helix coiled-coil rod homologue [Homo sapiens]		struct	52644507, 52646842, 35696286, 264092, 264094, 52645060, 35696052, 264107, 29331830, 52644045, 265006, 265007, 265008, 52644286, 52644228, 284668, 21905765, 21906766, 35695617, 265020, 52644150, 2653987, 33657109, 27486265, 35693763, 16108370, 263974, 16108374, 16108376, 52644332, 263961, 16108385, 264508, 264909, 264596
2370	78604120 (4739, 4740)			UNCLASSIFIED	264389
2371	57280408 (4741, 4742)			UNCLASSIFIED	263967, 263961
2372	67642413 (4743, 4744)			UNCLASSIFIED	29331826, 285010, 285016, 35669817, 264634, 60432113
2373	67416811 (4745, 4746)	Novel Protein sim. GBank gll4589562[db]BAA76613.11 - (AB023186) KIAA088 protein [Homo sapiens]		UNCLASSIFIED	265006, 265007, 265008, 265009, 265011, 264766, 35695917, 35695655, 263961, 284557, 264585
2374	94123685 (4747, 4748)	Novel Protein sim. GBank gll5105131[db]BAA60445.11 - (AP000081) 246aa long hypothetical ribonuclease PH [Aeropyrum pernix]	Contains protein domain (PF01136) - 3' exonuclease family	UNCLASSIFIED	60432049, 29331824, 264807, 52644045, 284512, 60433356, 21906754, 52644296, 67186559, 264446, 21906765, 21906766, 21606769, 33657023, 16106368, 55611576, 52644332
2375	67731355 (4749, 4750)	Novel Protein sim. GBank gll1351115[db]P47758[SRP_MOUSE - SIGNAL RECOGNITION PARTICLE RECEPTOR BETA SUBUNIT (SR-BETA)]		UNCLASSIFIED	264256, 26331630, 264909, 264910, 265009, 60433436, 21906754, 265017, 265018, 265019, 264682, 264266, 264665, 21906767, 263972, 3569655, 67168518, 60432113
2376	67613744 (4751, 4752)	Novel Protein sim. GBank gll2645435 (AF007780) - CHD3 [Drosophila melanogaster]	Contains protein domain (PF00626) - ATPase-associated PrD-finger	UNCLASSIFIED	16106394, 65274572, 22278997, 22276999, 264095, 29331822, 26147620, 29331824, 66714117, 29331825, 29331826, 29331826, 33658970, 29146498, 29146469, 264509, 265006, 265007, 265008, 265009, 60170631, 265010, 285011, 265018, 55611150, 16106351, 264764, 264288, 21906767, 21906768, 26146027, 29146829, 265021, 33657023, 33657109, 16106370, 16106374, 16106376, 35696423, 284556, 63373044, 16106365, 16106388, 56526466, 22279000, 22278002, 264563
2377	95319669 (4753, 4754)	Novel Protein sim. GBank gll5257005[db]AAC41239.11 - (AF063249) Rb binding protein homolog [Homo sapiens]	Contains protein domain (PF01388) - ARID DNA binding domain	UNCLASSIFIED	65274572, 56162575, 35696266, 264259, 29331822, 26331624, 66714117, 29331625, 26331626, 60432268, 29331827, 56162435, 284510, 265009, 60433356, 87166474, 285011, 265016, 264288, 21906765, 33657023, 264557, 56162323, 63373044, 16106365, 22278002, 264462
2378	94137032 (4755, 4756)	Novel Protein sim. GBank gll1072196 (U40942) - No definition line found [Caenorhabditis elegans]		UNCLASSIFIED	265017, 264266, 21906766
2379	65444324 (4757, 4758)	Novel Protein sim. GBank gll33337357 (AC004461) - hypothetical protein [Arabidopsis thaliana]	Contains protein domain (PF00400) - WD domain, G-beta repeat	Kinase	

2380	88923062 (4758, 4760)	Novel Protein sim. GBank gll4502839pefNP_001645.1PCOL1 - collagen, type XI, alpha 1	Contains protein domain (PF01410) - Fibrillar collagen C-terminal domain	collagen	264908, 264810, 265011
2361	67608241 (4761, 4762)	Novel Protein sim. GBank gll4455609emh(CAB3555) - (AL031846) dJ742C18.5 (novel Chromobox protein) [Homo sapiens]	Contains protein domain (PF00385) - Chromo (Chromatin Organization MOfifier) domain	helicase	5616257, 264091, 264083, 264259, 29331825, 264105, 264096, 60433356, 21906754, 265017, 265018, 264683, 264286, 264865, 264688, 264687, 264691, 264692, 264893, 55811576, 264638, 264567
2362	91225962 (4763, 4764)	Novel Protein sim. GBank gll4325130gplAAD1726f - (AF119716) dM1-2 protein [Drosophila melanogaster]	Contains protein domain (PF00828) - PHD-finger	transport	29331824, 60432289, 264805, 264596, 21906754, 264788, 265022, 264683, 263987, 33657109, 264629, 264831, 264556, 83373044, 60432113, 264482
2383	67442841 (4765, 4766)	Novel Protein sim. GBank gll19029621dbjBAA19005f - (DB9048) lectin-like oxidized LDL receptor [Bos taurus]	Contains protein domain (PF00389) - Lectin C-type domain	glycoprotein	265009, 21906765, 21906786
2384	95354766 (4767, 4768)	Novel Protein sim. GBank gll246285f (AF016252) - Spinophilin [Rattus norvegicus]	Contains protein domain (PF00589) - PDZ domain (Also known as DHR or GLGF)	struct	264488, 52644507, 52645156, 52646365, 35896286, 22276998, 52645080, 29331824, 29331826, 35896052, 29331828, 264806, 264828, 52644045, 265006, 265008, 265009, 33108954, 33657084, 52644296, 265011, 265017, 265016, 264683, 52644229, 21906785, 21906787, 21908768, 285020, 52644150, 33657023, 264683, 65274620, 52645128, 33657109, 33657182, 27486261, 27486282, 27488264, 33657349, 27486265, 35685763, 16108374, 35695855, 264634, 264555, 264556, 264557, 52644332, 264556, 264559, 18108365, 22279000, 22279002, 264488, 52644507, 52645156, 264687, 52646365, 22278995, 22278998, 22278997, 22278999, 264259, 52645080, 29331822, 29331828, 35896052, 52644045, 265006, 265007, 265008, 265009, 264910, 80432229, 60433356, 52646317, 21906754, 285019, 264446, 264683, 264686, 264687, 264689, 21906765, 21906766, 21906767, 21908789, 55611957, 265021, 265022, 264690, 264691, 264692, 65274620, 33657109, 16108370, 264631, 52644332, 22279000, 22279002, 264563, 264585, 264567
2385	95419485 (4769, 4770)			UNCLASSIFIED	
2386	94742849 (4771, 4772)	Novel Protein sim. GBank gll4929699gplAAD34110.1(AFI5187 - (AF151873) CGI-115 protein [Homo sapiens])		glycoprotein	264488, 22278985, 22278986, 22278987, 264259, 29146496, 264112, 264511, 60170831, 60432229, 264595, 60433438, 67166474, 67166559, 264682, 21906765, 21906766, 21906767, 21906769, 29146629, 35685817, 265021, 264680, 33657109, 264628, 16108376, 83373044, 60432113, 22279000, 264584, 264566, 264467

2367	14987990 (4773, 4774)			UNCLASSIFIED	264634
2368	11424604 (4775, 4776)			UNCLASSIFIED	264595
2389	95310650 (4777, 4778)	Novel Protein sim. GBank gl4759056[refNP_004372.1]pCREB - CAMP responsive element binding protein-like 1	Contains protein domain (PF00170) - bZIP transcription factor	dna_ma_bind	264488, 22276998, 22276999, 264509, 264805, 264908, 264907, 264908, 264909, 265006, 264511, 264512, 264910, 264591, 21908754, 264601, 264604, 264761, 18106351, 264784, 264266, 264786, 264786, 264788, 21906765, 21906766, 264892, 264693, 35698423, 264635, 264636, 264555, 63373044, 22279000, 264488
2390	94320912 (4779, 4780)	Novel Protein sim. GBank gl1644239[db]BAA122231 - (DB4103) mitochondrial DNA polymerase gamma [Homo sapiens]	Contains protein domain (PF00476) - DNA polymerase family A	polymerase	52644507, 56182575, 22278995, 35696286, 22276996, 22276997, 22276998, 26331822, 28331825, 28331826, 35696052, 264905, 52644045, 265009, 264758, 264759, 33108954, 52644296, 65658542, 265011, 265017, 265016, 264605, 52644229, 21906765, 21906767, 21906766, 21906769, 35695917, 52644150, 33657023, 33657108, 33657349, 35695783, 16108370, 16108374, 16108376, 35698423, 35695855, 264555, 52644332, 56182323, 60170394, 63373044, 56526488
2391	60036194 (4781, 4782)			UNCLASSIFIED	263976
2392	94245018 (4783, 4784)	Novel Protein sim. GBank gl16424016[db]BAA74863.1 - (AB020647) KIAA0840 protein [Homo sapiens]	Contains protein domain (PF00560) - Leucine Rich Repeat	nuclease	35696286, 35696052, 264508, 264905, 264509, 264908, 264907, 264908, 264909, 264510, 264511, 264512, 264910, 265009, 264591, 264758, 264800, 264804, 264782, 264448, 264784, 264389, 264768, 264768, 264769, 264688, 35695917, 264628, 16108374, 263976, 35698423, 35695855, 264631, 264634, 264635, 264636, 264637, 264638, 60170394, 264639, 264565, 264466

2383	95302633 (4785, 4788)	Novel Protein sim. GBank gi4506867 ref NP_000933.1 pRPLP - ribosomal protein, large, P0	Contains protein domain (PF00466) - ribosomal prot Ribosomal protein L10		181083392, 60424178, 264489, 18108394, 18108397, 22278995, 56894075, 35686286, 22278996, 22278997, 22278998, 264093, 60432049, 264258, 29331822, 29147620, 20281099, 29331824, 29331825, 68714117, 60432228, 29331826, 29331827, 29331828, 35696052, 29146499, 264508, 264509, 264905, 264907, 264908, 66712502, 52644045, 264828, 264808, 56182435, 264112, 264113, 264510, 265006, 264511, 265007, 265008, 265009, 264810, 264591, 264593, 60433356, 264595, 60433436, 52646317, 33109854, 21906754, 55811386, 265010, 265011, 265017, 265018, 265019, 264681, 264762, 18108351, 264763, 264682, 264764, 264683, 264369, 264266, 18108354, 264766, 264686, 264687, 264688, 264689, 18108359, 21906765, 21906766, 21906767, 21906768, 21906769, 55611957, 29148628, 29148784, 35685917, 265021, 265022, 33657023, 264692, 264693, 18108384, 33657109, 18108388, 27486261, 27486262, 33657349, 35685783, 18108370, 263972, 264628, 18108374, 263977, 18108378, 263978, 55810784, 35686423, 35685855, 264834, 60431850, 264555, 264637, 264557, 263981, 264558, 18106391, 60170394, 35686286, 22278997, 22278998, 58182181, 35686052, 265006, 264592, 55811386, 285010, 265011, 265017, 265019, 264448, 264883, 264288, 21906765, 21906768, 21908769, 55811957, 35689917, 33657023, 65274620, 33657182, 33657349, 35685763, 18108374, 18108378, 55810764, 55811576, 35686423, 60170394, 18108385, 264564, 264566, 264567
2384	94323266 (4787, 4788)	Novel Protein sim. GBank gi45159888 (AC004908) - zinc finger protein from gene of uncertain exon structure, similar to O98676 (P1D3025333) [Homo sapiens]	Contains protein domain (PF00096) - zinc finger, C2H2 type	dna_ma_bind	264259, 29331824, 264910, 264288, 265021, 83373044, 18108387, 264563, 264568
2385	95267212 (4789, 4790)	Novel Protein sim. GBank gi45712756 p AAC47638.1 AF16079 - (AF180798) calcium transporter Cat1 [Rattus norvegicus]			

2398	95086700 (4791, 4792)	Novel Protein sim. GBank g1105322[p11]834087 - hypothetical protein (L14-3' region) - human	Contains protein domain (PF00560) - Leucine Rich Repeat	nuclease	52646395, 16108397, 56182575, 35696286, 22276997, 22276999, 60432049, 264259, 29331822, 29331824, 29331825, 29331826, 29331827, 29331828, 264108, 264907, 29331830, 66712502, 264110, 60170631, 264591, 33857402, 60433438, 55812038, 33109954, 21806754, 33657084, 67168474, 265017, 265018, 265019, 264780, 264446, 264268, 264768, 52844229, 21906766, 21906767, 265020, 265021, 60170615, 264692, 33857023, 65274620, 52845129, 33657182, 27486262, 27466264, 27486265, 264629, 18108374, 35698423, 35698655, 264831, 264556, 52844332, 264558, 83373044, 18108386, 87168518, 22279002, 264482
2397	87260854 (4793, 4794)	.			52844507, 52645156, 56182575, 264259, 29147620, 264905, 264907, 264908, 264909, 264910, 264758, 52644296, 264603, 264804, 264762, 264681, 264784, 18108357, 264789, 21906768, 264693, 264628, 264635, 264638, 264639, 264564
2396	88047869 (4795, 4796)	Novel Protein sim. GBank g112558609 (AC005178) - H53 GST [Homo sapiens]		UNCLASSIFIED	
2399	87736965 (4797, 4798)	Novel Protein sim. GBank g11786117 (L41834) - nuclear protein [Ensis minor]		UNCLASSIFIED	56894075, 22276997, 22276998, 22276999, 264259, 29331822, 29331824, 66714117, 60432289, 264908, 29331830, 56182435, 264112, 264910, 33108954, 21806754, 87168474, 264600, 265017, 265018, 265019, 264784, 264765, 21906765, 21906766, 21906767, 21906769, 35695917, 265020, 265022, 60170615, 33657023, 18108370, 18108374, 264556, 60170394, 264558, 87168518, 22279000, 22279002, 264564, 264566, 264487
2400	81214116 (4799, 4800)	Novel Protein sim. GBank g11235822[g1]A869285.1 - (AF008945) glucose-6-phosphatase [Haplochromis nubilus]		phosphatase	21906766, 52646842, 56894075, 33657182, 27486262, 52644286, 265017

2401	91214118 (4801, 4802)	Novel Protein sim. GBank g12352822[gblpAAB69285.1 - (AF008845) glucose-6-phosphatase [Heplocitronia nudius)]		phosphatase	52844507, 52845156, 52844220, 284886, 21908784, 21908785, 52848385, 52848842, 21908786, 21908787, 21908788, 22278995, 35895917, 56894075, 35896286, 22278996, 22278997, 285020, 22278998, 22278999, 264259, 33657023, 52845080, 284693, 29331824, 33657109, 52845129, 29331826, 33657182, 28331827, 35896052, 27486281, 27486282, 33858870, 33657349, 27486285, 35895855, 285008, 285007, 265008, 265009, 264637, 52844332, 55812038, 52846317, 18108385, 52844298, 87168474, 285010, 87168559, 80432113, 265017, 265018, 265019, 264563, 264288, 264907, 264908, 284809, 284566
2402	91221406 (4803, 4804)	Novel Protein sim. GBank g14689258[gblpAAD27832.1(AF12185 - (AF121859) sorting nexin 9 [Homo sapiens]			
2403	84135432 (4805, 4806)	Novel Protein sim. GBank g14629575[gblpAAD34046.1(AF15181 - (AF151811) CGI-53 protein [Homo sapiens]	Contains protein domain (PF00062) - C-type lysozyme/alpha-lactalbumin family		22278999, 35898052, 265018, 264886, 264693, 83373044, 264567
2404	85312805 (4807, 4808)	Novel Protein sim. GBank g12315786 (AF016685) - similar to short chain-type dehydrogenases [Caenorhabditis elegans]	Contains protein domain (PF00106) - short chain dehydrogenase	dehydrogenase	35898286, 28331826, 35896052, 265008, 265018, 21908789, 264564
2405	84311851 (4809, 4810)	Novel Protein sim. GBank g14641781[gblpBAA035811 - (D14853) polyprotein [Hepatitis C virus]		UNCLASSIFIED	35896286, 29331822, 285007, 21908744, 285017, 265018, 285019, 284783, 264368, 21908785, 35895917, 285020, 265021, 52844150, 284683, 35895855, 264632, 52844332, 22279002
2406	88094501 (4811, 4812)	Novel Protein sim. GBank g12773363 (AF041382) - microtubule binding protein D-CIP-190 [Drosophila melanogaster]	Contains protein domain (PF01302) - CAP-Gly domain	strud	52848842, 22278994, 22278996, 35896288, 22278997, 22278998, 22278999, 284092, 264083, 60432049, 264259, 29331822, 29331824, 29331825, 29331827, 35898052, 29331828, 264102, 264106, 264908, 52844045, 285007, 285008, 285009, 284910, 264582, 80433356, 60433438, 33109954, 265010, 265011, 265018, 265019, 264369, 284885, 284686, 21908788, 21908789, 52844150, 264693, 52845129, 264628, 35896423, 264632, 56182323, 264638, 22279000, 22279002, 264583
2407	78465005 (4813, 4814)	Novel Protein sim. GBank g1423442[pil]S33513 - gene F11 protein - mouse		UNCLASSIFIED	284685, 264886
2408	87381503 (4815, 4816)	Novel Protein sim. GBank g1423442[pil]S33513 - gene F11 protein - mouse		UNCLASSIFIED	264910, 285010, 264448, 264557

2408	66741770 (4817, 4818)	Novel Protein sim. GBank g 1178801 sp P45969 N26, CAEEI - HYPOTHETICAL 20.8 KD PROTEIN T08A5.8 IN CHROMOSOME III		UNCLASSIFIED	22278985, 22278986, 22278987, 264087, 29331822, 29331824, 29331827, 29146498, 52644045, 60433438, 33657084, 87168474, 264780, 21906767, 29148627, 29148629, 52644150, 33657023, 263967, 20281069, 18108374, 20281071, 56182323, 83373044, 18108385, 87168518 22278989, 264259
2410	87804860 (4819, 4820)	Novel Protein sim. GBank g 14968262 gb AAC48052.2 - (UBA649) Contains similarity to Plam domain: PF00646 (F- box), Score=28.7, E-value=4.3e-05, N=1 [Caenorhabditis elegans]	Contains protein domain (PF00080) - Copper/zinc superoxide dismutase (SODC)		
2411	87534833 (4821, 4822)	Novel Protein sim. GBank g 13114713 (AF061346) - Ebp1 protein [Mus musculus]		inf	29331824, 29331827, 29331828, 264764, 264368, 33657109, 56182323 22278989, 29331827, 264807, 265011, 265017, 265018, 265019, 18108351, 21906768, 285020, 33657109, 264559, 18108385
2412	87778332 (4823, 4824)	Novel Protein sim. GBank g 15410338 gb AAD43038.1 - (AF106885) myelin gene expression factor 2 [Homo sapiens]	Contains protein domain (PF00070) - RNA recognition motif, (a.k.a. RRM, RBD, or RNP domain)	dna_rna_bind	264488, 264259, 29331826, 264508, 264905, 264509, 264908, 264907, 264908, 264510, 264511, 264512, 265008, 265009, 264910, 285011, 264682, 264784, 264768, 264888, 264788, 264889, 265021, 33657023, 18108370, 264628, 33659485, 264832, 264834, 264835, 264838, 83373044, 264563, 264564, 264585, 264586, 264587, 264486 52645158, 52646842, 52646385, 18108398, 56182575, 22278994, 22278995, 56994075, 22278986, 33696286, 22278987, 22278988, 22278989, 264259, 52645080, 29331822, 29331824, 29331825, 29331828, 29331827, 29331828, 33696052, 33656970, 29331830, 264909, 60433358, 33657402, 264594, 52646317, 21906754, 33657084, 265010, 87168559, 285017, 285018, 265019, 264369, 264684, 264687, 264888, 56181562, 21906764, 264889, 21906785, 21906788, 21906787, 29148627, 21906789, 285020, 265021, 60170615, 33657023, 264893, 52645128, 33657108, 33657182, 27486261, 27486262, 27486285, 33657348, 18108370, 60431528, 264828, 18108374, 18108378, 55610764, 264638, 52644332, 264638, 264558, 56182323, 83373044, 18108385, 87168518, 22279002
2413	94133820 (4825, 4828)	Novel Protein sim. GBank g 15262705 emb CAB45778.1 - (AL080214) hypothetical protein [Homo sapiens]	Contains protein domain (PF00038) - Intermediate filament proteins	struct	
2414	84312580 (4827, 4828)	Novel Protein sim. GBank g 1082340 pf S52863 - DNA- binding protein R kappa B - human		ubiquitin	

2415	98089002 (4629, 4630)	Novel Protein sim. GBank g 423915 p K45439 - myosin I heavy chain - rat	Contains protein domain (PF00063) - Myosin head (motor domain)	struct	284259, 264908, 60433358, 33657402, 21906754, 285018, 264687, 264889, 21906789, 55611957, 285021, 264690, 264691, 33657023, 264693, 35698423, 56182323, 56526486
2416	94118356 (4631, 4632)	Novel Protein sim. GBank g 3025445 (AC004528) - R32164_1 [Homo sapiens]			264638
2417	87733334 (4633, 4634)	Novel Protein sim. GBank g 1084944 p I S54495 - hypothetical protein YPR021c - yeast (Saccharomyces cerevisiae)	Contains protein domain (PF00153) - Mitochondrial carrier proteins		264094, 29331822, 29331824, 29331827, 284389
2418	94234349 (4635, 4636)	Novel Protein sim. GBank g 1176572 p P45895 NAA_CAEEL - HYPOTHETICAL 91.0 KD PROTEIN PAR2.4 IN CHROMOSOME III	Contains protein domain (PF00411) - Ribosomal protein S11	UNCLASSIFIED	56994075, 264091, 264256, 29331824, 29331825, 60432289, 29331828, 264905, 264907, 264511, 265009, 60432229, 21906754, 87168559, 265019, 264682, 21906768, 21906769, 265020, 285021, 33657023, 85274620, 18108370, 55811576, 284634, 60170394, 16108365, 22279000, 22279002, 264568
2419	82374248 (4637, 4638)	Novel Protein sim. GBank g 284008 p I S16732 - autoantigen, G4K - human		struct	284569, 264782, 264448, 264691, 264631, 264634, 264555, 264556, 264838, 264556, 29331824, 29331825, 29331826, 60432229, 33109854, 85658542, 87188474, 265016, 264288, 265020, 264564
2420	94844244 (4639, 4640)	Novel Protein sim. GBank g 1076211 p I S50755 - hypothetical protein VSP-3 - Chlamydomonas reinhardtii		UNCLASSIFIED	264909, 264768, 264838
2421	87805345 (4641, 4642)	Novel Protein sim. GBank g 2224567 p BAA20772 - (AB00231) KIAA0313 [Homo sapiens]	Contains protein domain (PF00617) - RasGEF domain	UNCLASSIFIED	18108392, 18108394, 18108396, 264906, 265005, 265010, 18108351, 18108374, 18108385
2423	86056390 (4645, 4646)	Novel Protein sim. GBank g 4505153 p NP_002392.1 MEKK - MAP/ERK kinase 3	Contains protein domain (PF00069) - Eukaryotic protein kinase domain	Kinase	264259, 60432049, 29331822, 29331826, 60432289, 29331828, 265008, 265009, 60433356, 21906754, 265017, 265018, 265019, 21906766, 21906768, 21908769, 285020, 285021, 20281149, 263971, 60432113
2424	94854047 (4647, 4648)	Novel Protein sim. GBank g 2988398 (AC004381) - Unknown gene product [Homo sapiens]		UNCLASSIFIED	56182575, 35896286, 22278997, 60432048, 264259, 29331828, 29331828, 264905, 66712502, 29331830, 60433356, 265011, 285019, 284766, 21906768, 55611957, 284892, 33657023, 33657109, 55811576, 56182323, 63373044, 16108385, 18108386, 60432113, 22279000
2425	87415981 (4649, 4650)	Novel Protein sim. GBank g 2077832 p BAA19879 - (DB8558) Protein Kinase [Rattus norvegicus]	Contains protein domain (PF00069) - Eukaryotic protein kinase domain	kinase	264834
2426	87613945 (4651, 4652)	Novel Protein sim. GBank g 2039368 p AA53003.1 - (UB4619) circulating cathodic antigen [Schistosoma mansoni]		UNCLASSIFIED	22278996, 22278996, 264258, 264102, 264512, 265008, 21908767, 18108370, 18108374, 263976

2427	87622893 (4853, 4854)	Novel Protein sim. GBank gll4680895jgblAAD2737.1(A.F13296 - (A.F132962) CGI-28 protein [Homo sapiens]	Contains protein domain (PF00573) - Ribosomal protein L4/L1 family	ribosomeprot	264258, 20261098, 35696052, 265006, 264594, 265011, 264780, 18108331, 264682, 264863, 264369, 264684, 264688, 264687, 264689, 21906766, 264691, 264692, 18106374, 16108377, 264557, 264639, 18108335
2428	85732888 (4855, 4856)	Novel Protein sim. GBank gll1537070 (U63840) - nucleoporin p54 [Rattus norvegicus]			22278996, 22278998, 35696052, 21906754, 264288, 21906765, 21906768, 21906769, 35695917, 265020, 263972, 22278902
2429	87769276 (4857, 4858)	Novel Protein sim. GBank gll801931 (M94316) - neurofilament-H [Oryzodagus curvicolus]	Contains protein domain (PF00711) - Beta defensins	UNCLASSIFIED	22278993, 29331824, 264806, 264909, 264511, 265009, 21906754, 285017, 265016, 265019, 264448, 264883, 264288, 21906765, 21906768, 265021, 264693, 18108381
2430	86948827 (4859, 4860)	Novel Protein sim. GBank gll3860729lemb(CAA14630) - (A.J235270) CELL DIVISION PROTEIN FTSJ (fts.) [Rickettsia prowazekii]	Contains protein domain (PF01728) - FtsJ cell division protein	UNCLASSIFIED	264112, 264681
2431	87848884 (4861, 4862)	Novel Protein sim. GBank gll3876367lemb(CA83287) - (Z69360) Weak similarity to Eimeria thrombospondin (PIR Acc. No. A45517): cDNA EST EMBL.M89268 comes from this gene: cDNA EST yk285b8.5 comes from this gene [Caenorhabditis elegans]	protease		29331828, 29331627, 35696052, 2914849, 264905, 264806, 264681, 264286, 264689, 21906765, 264692, 35696423
2432	80083033 (4863, 4864)	Novel Protein sim. GBank gll2224593jdbjBAA207841 - (AB002324) KIAA0328 [Homo sapiens]	Contains protein domain (PF00096) - Zinc finger, C2H2 type		264634, 264556
2433	80055092 (4865, 4866)	Novel Protein sim. GBank gll1283289 (U47856) - fibrin-4 [Acanthamoeba dioleumatus]		transcription factor	264589, 284805, 265018, 284762, 264683, 264891, 284558, 284557, 284638, 264558
2434	19520148 (4867, 4868)	Novel Protein sim. GBank gll13641352 (AF091234) - putative transcription factor [Mus musculus]		UNCLASSIFIED	264555
2435	20758044 (4869, 4870)	Novel Protein sim. GBank gll3641352 (AF091234) - putative transcription factor [Mus musculus]		UNCLASSIFIED	265008, 264756, 265010, 264688, 27486261, 263972, 18108374, 18106381
2436	68044008 (4871, 4872)	Novel Protein sim. GBank gll3641352 (AF091234) - putative transcription factor [Mus musculus]		UNCLASSIFIED	29331828, 265007, 265008, 265017, 284760, 264685, 264693, 264565
2437	63363424 (4873, 4874)	Novel Protein sim. GBank gll3860014 (AF091088) - unknown [Homo sapiens]	Contains protein domain (PF01256) - Unclassified protein family	UNCLASSIFIED	29331826, 264508, 264805, 264508, 264808, 264907, 264908, 264909, 264511, 265008, 264512, 264910, 265008, 264591, 33857402, 21906754, 265011, 264760, 264764, 264685, 264688, 264766, 35695917, 33657023, 264693, 264631, 264632, 56182323, 264558, 63373044, 264563, 264564, 264565, 264568, 264587

2439	94850850 (4877, 4878)	Novel Protein sim. GBank gij263518jbuAA0153451 - (AC004044) small nuclear riboprotein Sm-D1 [Arabidopsis thaliana]	Contains protein domain (PF01423) - UNCLASSIFIED	60424178, 18108397, 56182575, 56186166, 56964075, 22278998, 35898286, 22278997, 22278999, 264259, 52845080, 29331822, 56182181, 29331824, 80424269, 66714117, 29331825, 60432288, 29331826, 29331827, 29331828, 35898052, 29146498, 264509, 264905, 264906, 52844045, 80431735, 33109954, 21908754, 33657084, 55811388, 52644296, 87188474, 265017, 265018, 265019, 18108351, 264448, 284288, 264768, 52844229, 56181562, 21908764, 21906765, 21906766, 21906787, 21906768, 21906769, 35895917, 33657023, 33657109, 33657182, 27486262, 27486264, 33657349, 27486265, 35895763, 18108370, 80431528, 263977, 55810764, 35696423, 85274791, 35695885, 60431850, 58182323, 60432113, 22279000, 22279002, 264567
2440	87841733 (4878, 4880)		UNCLASSIFIED	29331822, 29331824, 29331825, 29331827, 35896052, 284508, 264807, 264510, 265018, 265019, 264448, 264369, 265020, 265021, 58182323, 284639, 22278002
2441	87823914 (4881, 4882)	Novel Protein sim. GBank gij3024869ispj56524ly288, HUMAN - HYPOTHETICAL PROTEIN KIAA0288 (HA8118)	UNCLASSIFIED	284486, 284629, 18108374, 264564
2442	87273590 (4883, 4884)	Novel Protein sim. GBank gij4506013jrefjnp_002703, 1jppPP1 - protein phosphatase 1, regulatory subunit 7	Contains protein domain (PF00360) - UNCLASSIFIED Leucine Rich Repeat	22278996, 22278999, 29331822, 264768, 264693
2443	94305949 (4885, 4888)	Novel Protein sim. GBank gij1170658jsepjC02973jKID1, RAT - RENAL TRANSCRIPTION FACTOR KID-1 (TRANSCRIPTION FACTOR 17)	Contains protein domain (PF01352) - TRANSCRIPTION FACTOR	264808
2444	88086345 (4887, 4888)	Novel Protein sim. GBank gij4758824jrefjnp_004280, 1jppNF3 - nuclear factor (erythroid-derived 2)-like 3	Contains protein domain (PF00170) - TRANSCRIPTION FACTOR	264259, 18108382, 18108383, 18108385, 22279000
2445	87338638 (4889, 4890)	Novel Protein sim. GBank gij2133950jprjSS6222 - P.Q-rich protein - human	bZIP transcription factor	264259, 35696052, 264369, 18108381
2446	88058293 (4891, 4892)	Novel Protein sim. GBank gij4753887jtembJCA05409.21 - (A1002424) p85 protein [Rattus norvegicus]	Contains protein domain (PF00095) - proteaseinhib	265011, 264689, 33657023, 263961, 18108385
2447	94845149 (4893, 4894)	Novel Protein sim. GBank gij4685813jrefjnp_005408, 1jppST5j - suppression of tumorigenicity 5	WAP-type (Whey Acidic Protein) 'four-disulfide core'	58182575, 264259, 29331824, 29331825, 29331827, 60433356, 60433438, 264758, 265018, 264692, 85274620, 60431528, 85274791, 58182323

2448	67749860 (4895, 4896)			UNCLASSIFIED	22276998, 22276997, 22276998, 29331826, 35696052, 264107, 264110, 67166474, 87166559, 16106351, 21906767, 21906769, 27466282, 263976
2449	8769075 (4897, 4898)	Novel Protein sim. GBank gij728837[ispP39194/ALUT_HUMAN - III; ALU SUBFAMILY SQ WARNING ENTRY III]		cadherin	264259, 264628, 265007, 264595, 265021, 56526468
2450	86597784 (4898, 4900)			UNCLASSIFIED	264906
2451	91014563 (4901, 4902)	Novel Protein sim. GBank gij1710021[ispP35290/RB24_MOUSE - RAS-RELATED PROTEIN RAB-24 (RAB-16)]	Contains protein domain (PF00071) - Ras family	glycoprotein	264083, 29331822, 29331824, 29331825, 66714117, 29331828, 29331828, 35698052, 264907, 66712502, 29331830, 264910, 265009, 264758, 265017, 265018, 264762, 264448, 264288, 21906767, 265021, 33657023, 264693, 33657109, 263969, 63373044, 16106385
2452	81230508 (4903, 4904)	Novel Protein sim. GBank gij1504034[dpj]BA132181 - (D86980) KIAA0227 [Homo sapiens]		isomerase	264102, 264112, 264668, 263972, 18108374, 63373044, 264563
2453	84201068 (4905, 4906)	Novel Protein sim. GBank gij72680079 (AC004142) - similar to murine leucine-rich repeat protein; possible role in neural development by protein-protein interactions; 93% similarity to D48802 (PIDD; g1369908) [Homo sapiens]	Contains protein domain (PF00560) - Leucine Rich Repeat	ingrecep	264509, 264512, 18108385
2454	95310891 (4907, 4908)	Novel Protein sim. GBank gij1078802[pvr]S49915 - extensin like protein - maize	Contains protein domain (PF00170) - bZIP transcription factor	UNCLASSIFIED	263994, 66714117, 29331827, 264508, 264509, 264905, 264906, 264907, 264908, 264909, 264510, 264511, 264512, 265009, 264910, 264591, 264758, 264759, 265010, 265011, 264603, 264604, 264760, 264761, 264762, 18108351, 264764, 264765, 264766, 264688, 264768, 264769, 264534, 264691, 264692, 33657023, 264693, 33657109, 264628, 263976, 35698055, 264634, 264635, 264637, 264638, 264639, 83373044, 16106385, 264583, 264564, 264486
2455	95286301 (4909, 4910)	Novel Protein sim. GBank gij543817[ispP35655/AA47_MOUSE - CLATHRIN COAT ASSEMBLY PROTEIN AP47 (CLATHRIN COAT ASSOCIATED PROTEIN AP47) (GOLGI ADAPTOR AP-1 47 KD PROTEIN) (HA1 47 KD SUBUNIT) (CLATHRIN ASSEMBLY PROTEIN ASSEMBLY PROTEIN COMPLEX 1 MEDIUM CHAIN)]	Contains protein domain (PF00928) - glycoprotein Adaptor complexes medium subunit family		264488, 22276998, 264259, 35696052, 264905, 264906, 264907, 264908, 264909, 264510, 264512, 265008, 265009, 264910, 264591, 264592, 264593, 264594, 264758, 265018, 264760, 264681, 18108351, 264683, 264764, 18108354, 264766, 264768, 264769, 264688, 21906766, 21906767, 21906769, 29148629, 35695917, 285020, 285022, 33657023, 33657109, 18108370, 264628, 264629, 264631, 264632, 264635, 56182323, 60170394, 16106385, 264583, 264564, 264565, 264567
2456	88188700 (4911, 4912)	Novel Protein sim. GBank gij2586630 (AC003079) - Arkylrn like; 54% similar to 20223404 (NIDg1092123) in exons spanning 43974 to 11551 of clone. [Homo sapiens]	Contains protein domain (PF00023) - Ark repeat	Kinase	264693

2457	94118375 (4913, 4914)	Novel Protein sim. GBank gij3025447 (AC004528) - R32184_3 [Homo sapiens]		UNCLASSIFIED	56181686, 264805, 264907, 264511, 264596, 55811386, 264682, 264684, 264685, 264687, 264691, 33657023, 284693, 35695855, 264636, 264555, 58182323, 284558, 56526486, 264563, 264691, 264693, 264634, 264559
2458	85875304 (4915, 4918)	Novel Protein sim. GBank gij2364942 (AF022985) - Similar to collagen [Caenorhabditis elegans]		UNCLASSIFIED	27486265
2459	87551913 (4917, 4918)	Novel Protein sim. GBank gij5441942 (gb AA043187.1 AC00498) supported by mouse EST AA538043 (MID:g2284036) [Homo sapiens]	Contains protein domain (PF00097) - Zinc finger, C2HC4 type (RING finger)	transport	
2460	94315289 (4919, 4920)	Novel Protein sim. GBank gij4929701 (gb AA043111.1 AF15187) - (AF151874) CGI-116 protein [Homo sapiens]		kinase	65274572, 35696286, 22278996, 22278887, 60432049, 29331822, 29331624, 29331825, 66714117, 60432289, 29331826, 29331827, 29331828, 33656970, 29148499, 284102, 284109, 80433438, 265017, 265018, 285019, 284288, 21906765, 21906766, 21906769, 35695917, 285020, 284691, 33657023, 27486261, 18108374, 35695855, 87168518, 60432113
2481	87645147 (4921, 4922)	Novel Protein sim. GBank gij442662 (gb AA020633) - (AF126062) Act-like 2 binding protein BART1 [Homo sapiens]		UNCLASSIFIED	264258, 29331828, 264910, 18108351, 18108370, 18108374
2482	86988002 (4923, 4924)	Novel Protein sim. GBank gij5420367 (emb CAB46678.1 - (AJ243459) proteolipophoglycan [Lishmania major]			264809, 264756, 264684, 18108374, 264637, 18108385
2483	84386543 (4925, 4926)	Novel Protein sim. GBank gij5052518 (gb AA036588.1 AF14561) - (AF145613) BCDNA, GH03108 [Drosophila melanogaster]	Contains protein domain (PF00096) - Zinc finger, C2H2 type	UNCLASSIFIED	264681, 264566
2484	91218957 (4927, 4928)	Novel Protein sim. GBank gij5410300 (gb AA043021.1 - (AF100757) COP9 complex subunit 4 [Homo sapiens]	Contains protein domain (PF01398) - PCI domain	protease	284489, 52846842, 22278995, 35696286, 22278996, 22278997, 22278999, 264258, 29331822, 29331824, 66714117, 29331825, 29331826, 29331827, 35696052, 29331828, 284509, 52844045, 264510, 284511, 284512, 265008, 60170831, 264583, 52846317, 33109954, 33657084, 285017, 265018, 285019, 284782, 264448, 264784, 264288, 264766, 21906765, 21906768, 21906769, 21906769, 285021, 33657023, 33657108, 18108370, 18108381, 80170394, 18108385, 22279002, 284486

2465	95357483 (4928, 4930)	Novel Protein sim. GBank gll4508401reNIP_002871.1 pPAF1 - v-vai-1 murine leukemia viral oncogene homolog 1	Contains protein domain (PF00069) - Eukaryotic protein kinase domain	oncogene	16106392, 52644507, 52645156, 52648385, 22278994, 22278995, 35696286, 22278996, 22278998, 264259, 29331822, 29331824, 29331825, 60424269, 60432289, 29331827, 35696052, 29331828, 284807, 29331830, 52644045, 284909, 56182435, 264511, 265007, 265008, 265009, 264910, 33657402, 60433438, 55812038, 21906754, 33109954, 265010, 265011, 87168559, 284600, 285017, 265018, 285019, 18108351, 284369, 284288, 264685, 264767, 21908785, 21908767, 21908788, 55811957, 35895917, 265020, 265021, 80170815, 52644150, 33657023, 52845129, 33657109, 27488261, 27486264, 35895763, 284828, 283972, 18108374, 35695855, 264638, 284637, 60170394, 56526486, 87188518, 60432113, 264563, 284584, 284588, 264487, 264389
2466	85681388 (4931, 4932)	Novel Protein sim. GBank gll4321619 gb AAD15788.1 - (AF051088) seven transmembrane domain orphan receptor (Mus musculus)			
2467	88059485 (4933, 4934)	Novel Protein sim. GBank gll513300 (AC005595) - F16801_1, partial COS (Homo sapiens)		UNCLASSIFIED	56994075, 264808, 21908788, 33657023
2468	87814696 (4935, 4936)	Novel Protein sim. GBank gll2143455 p l58106 - gene DKR-N9 protein - mouse (fragment)	Contains protein domain (PF00400) - WD domain, G-beta repeat	kinase	29331824, 52644045, 265008, 264810, 265019, 21908785, 21908789, 265021, 264288, 264628
2469	86284397 (4937, 4938)	Novel Protein sim. GBank gll5420389 emb CAB46880.1 - (A234460) proteophosphopcyan (Leishmania major)		UNCLASSIFIED	264582, 264682, 264555, 264558, 264557, 264558, 264559, 18108385, 264482
2470	80223631 (4939, 4940)	Novel Protein sim. GBank gll5419882 emb CAB46424.1 - (AL088749) DKFz343G153 (Homo sapiens)		UNCLASSIFIED	65274572, 35696288, 29331827, 265007, 264592, 33109954, 265018, 265019, 284208, 55811957, 265020, 284893, 55811576, 56182323
2472	85060811 (4943, 4944)	Novel Protein sim. GBank gll4929747 gb AAD4134.1 AF15189 - (AF151897) CGI-139 protein (Homo sapiens)		UNCLASSIFIED	284082, 60432049, 29331825, 60433356, 265010, 265011, 18108351, 264764, 264288, 264692, 65274820, 18108370, 18108372, 18108374, 264634, 18108385
2473	85421508 (4945, 4946)	Novel Protein sim. GBank gll4539009 emb CAB39830.1 - (AL049461) putative protein (Arachidopsis thaliana)			80424178, 65274572, 22278999, 60424269, 29331828, 265008, 60433358, 60433438, 265010, 18108351, 264448, 264288, 264687, 264689, 265021, 284892, 65274820, 60431528, 65274791, 264556, 56182323, 80432113
2474	94315818 (4947, 4948)	Novel Protein sim. GBank gll3252827 (AC004382) - Unknown gene product (Homo sapiens)			65274572, 58894075, 284259, 29331828, 80170831, 265017, 265018, 265019, 284683, 284369, 265020, 284893, 284583, 284584

2475	94321693 (4949, 4950)	Novel Protein sim. GBank gij1216486 (U48652) - HT protein [Cricetus griseus]	Contains protein domain (PF00008) - EGF-like domain	lgf	264259, 29331822, 265006, 265007, 265010, 265011, 264448, 264288, 264389, 264685, 264888, 18108357, 264768, 18108382, 264893, 18108370, 18108374, 18108379, 35696423, 83373044, 18108383, 18108385, 264564, 264565, 264567
2476	94315818 (4951, 4952)	Novel Protein sim. GBank gij1232827 (AC004382) - Unknown gene product [Homo sapiens]		UNCLASSIFIED	264259, 60424269, 66714117, 264905, 265006, 264511, 265008, 265009, 264758, 265010, 265011, 18108351, 264681, 264369, 264288, 264889, 21906767, 265020, 18108374, 264639, 18108382, 83373044, 18108385, 87168518
2477	20718974 (4953, 4954)			UNCLASSIFIED	263978
2478	17659165 (4955, 4956)			UNCLASSIFIED	265017
2479	94314569 (4957, 4958)	Novel Protein sim. GBank gij1644232dijlBAA110821 - (DBTDB) N-WASP [Bos taurus]		Im7	56994075, 22278998, 21906754, 264682, 21906785
2480	95295805 (4959, 4960)	Novel Protein sim. GBank gij1644232dijlBAA110821 - (AB028989) KIAA1068 protein [Homo sapiens]		UNCLASSIFIED	264905, 284907, 264765
2481	94718461 (4961, 4962)			collagen	55274572, 56182575, 22278997, 264094, 264259, 29331822, 29331824, 68714117, 29331827, 35696052, 264508, 284905, 284906, 284907, 284908, 52844045, 264909, 56182435, 265008, 264910, 33657402, 55812038, 264758, 265010, 265011, 265017, 265018, 264760, 264762, 18108351, 264764, 264288, 284768, 284686, 264768, 21906768, 55811957, 265020, 284691, 264692, 264693, 264629, 55811576, 264830, 264634, 264635, 264638, 264637, 264556, 264558, 56182323, 83373044, 80432113, 22278002
2482	87383165 (4963, 4964)	Novel Protein sim. GBank gij123249pji[S28407] - guanine nucleotide-exchange activator CDC25 homolog - mouse		UNCLASSIFIED	29331822, 29331824, 29331825, 29331827, 264508, 264905, 264509, 264908, 264907, 264908, 264511, 264591, 264788, 264693, 264691, 264692, 264693, 264693, 264583
2483	67731563 (4965, 4966)			UNCLASSIFIED	264488, 22278995, 264093, 264095, 60432049, 60433356, 60433436, 264448, 264288, 263967, 18108370, 18108385, 18108388, 264482
2484	94187774 (4967, 4968)	Novel Protein sim. GBank gij1228631ipjP39188/ALU1_HUMAN - III ALU SUBFAMILY J WARNING ENTRY III		kinase	264563
2485	87788556 (4969, 4970)	Novel Protein sim. GBank gij1185397 (U25261) - SH3 domain binding protein [Rattus norvegicus]		UNCLASSIFIED	22278995, 22278996, 22278997, 22278999, 264259, 60432049, 29331824, 60432289, 29331827, 265007, 264910, 264593, 264600, 264803, 284604, 265019, 264448, 264288, 264695, 264896, 264769, 264889, 35695917, 265022, 264692, 264693, 58182323

2486	87746978 (4871, 4872)	Novel Protein sim. GBank gij2662187 dbj BAA237151 - (AB007803) KIAA0443 [Homo sapiens]				265017, 264555
2487	95343105 (4873, 4874)	Novel Protein sim. GBank gij464559 sp P35287 RB14_RAT RAS-RELATED PROTEIN RAB-14	Contains protein domain (PF00071) - RAS family		glycoprotein	22278998, 35696266, 22276997, 22276998, 22278998, 264092, 264258, 28331822, 35698052, 264108, 264905, 264807, 2933130, 264909, 265008, 264511, 265008, 265009, 60433438, 21806754, 33108954, 87168559, 265016, 284681, 264288, 264687, 21908785, 21806768, 21806767, 21806788, 21908789, 35695917, 265021, 265022, 264534, 33657023, 264692, 33657109, 263872, 16108377, 35696423, 35695855, 60170394, 16108385, 56526486, 22278000, 22278002, 264563, 264482, 264585, 20281169, 16108391
2488	87652451 (4875, 4876)				UNCLASSIFIED	264810, 264448, 264288, 264684, 264691, 264834
2489	82890565 (4877, 4878)	Novel Protein sim. GBank gij4688439 emb CAB43355.1 - (AL050253) hypothetical protein [Homo sapiens]	Contains protein domain (PF00084) - Sushi domain (SCR repeat)		complement/recept	264886, 264693, 55611578, 22279002
2490	88069809 (4879, 4880)	Novel Protein sim. GBank gij2588624 (ACD03083) - Rap2 interacting protein-like; similar to U73841 (PID:g1916016) [Homo sapiens]			UNCLASSIFIED	264907, 265008, 22278002
2481	81242118 (4881, 4882)	Novel Protein sim. GBank gij726832 sp P39189 ALU2_HUMAN - III ALU SUBFAMILY SE WARNING ENTRY III			Im7	264258, 29331826, 265006, 264762, 16108370, 16108376, 16108379
2482	95308202 (4883, 4884)	Novel Protein sim. GBank gij3355303 (AF001549) - Unknown gene product [Homo sapiens]			transcript/lacior	264488, 22278998, 22278999, 29331826, 264591, 33108954, 265017, 55611550, 21806764, 21806768, 284692, 60431528, 87168516, 60432113, 22278000

2483	95422415 (4895, 4896)	Novel Protein sim. GBank g1420307 (dbj)DA14932.1 - (AB020718) KIAA0809 protein [Homo sapiens]	Contains protein domain (PF01424) - R3H domain	struct	16108394, 264867, 65274572, 56182575, 22278995, 56994075, 80432049, 29331822, 29331824, 29331825, 29331826, 29331827, 29146498, 264508, 264509, 264508, 264509, 264508, 264509, 264507, 265007, 264512, 265008, 265009, 264511, 265007, 264512, 265008, 265009, 264910, 21908754, 265011, 264600, 265017, 265018, 264604, 264805, 265019, 55811150, 264782, 18108351, 264881, 264448, 264883, 264369, 264288, 18108355, 18108357, 264887, 21908765, 21908768, 21908787, 21908788, 21908789, 265020, 264681, 264692, 33857023, 33657349, 18108370, 18108374, 18108376, 55810764, 18108378, 65274791, 264630, 264832, 264634, 264635, 264636, 264555, 264637, 264557, 264558, 264639, 264559, 83373044, 18108385, 87168516, 60432113, 22279000, 22279002, 264482, 264588, 264488
2484	30783118 (4897, 4898)	Novel Protein sim. GBank g15420389 (emb)CAB4680.11 - (AJ243480) proteophosphoglycan [Leishmania major]		UNCLASSIFIED	264807, 264601
2485	84234551 (4899, 4900)	Novel Protein sim. GBank g15420389 (emb)CAB4680.11 - (AJ243480) proteophosphoglycan [Leishmania major]		collagen	283894, 22278997, 33686052, 264509, 264805, 264908, 264807, 264808, 264809, 285006, 265008, 264595, 264604, 264448, 264682, 264784, 264288, 264685, 264786, 264789, 264689, 265020, 264692, 65274620, 264829, 55810764, 35698423, 55811578, 264838, 264637, 18108385, 22279000, 264564, 264587, 264488
2486	80018765 (4991, 4992)	Novel Protein sim. GBank g14808220 (emb)CAB42832.11 - (AL022315) d117715.1 (PUTATIVE novel protein) [Homo sapiens]		struct	29147620, 264805, 265006, 265007, 18108348, 18108382, 18108370, 18108374, 264555, 264556, 18108381, 18108383, 18108388
2487	91723554 (4993, 4994)			UNCLASSIFIED	52644507, 22278996, 22278999, 29331824, 29331828, 33857402, 21908754, 87186474, 265019, 264389, 264889, 21908785, 21908786, 21908787, 21908788, 265020, 33857023, 18108376, 18108387
2488	87724833 (4995, 4996)	Novel Protein sim. GBank g11200503 (U47924) - B [Homo sapiens]		UNCLASSIFIED	29331827, 264512, 264810, 264288, 18108374, 35698385
2489	84685125 (4997, 4998)	Novel Protein sim. GBank g13510234 (AC005581) - R31237.1, partial CDS [Homo sapiens]	Contains protein domain (PF00089) - Eukaryotic protein kinase domain	kinase	264808, 55812038, 264631, 264637, 264558

2500	84648324 (4898, 5000)	Novel Protein sim. GBank gi136812751emb CAA217251 - (AL032855) predicted using GeneFinder; similar to Inositol monophosphatase family; cDNA EST YK255e11.5 comes from this gene [Caenorhabditis elegans]	Inositol monophosphatase family	transport	52844507, 52845156, 22278995, 56994075, 35696286, 22276998, 284259, 52645080, 29331824, 29331825, 66714117, 60432289, 29331826, 29331827, 35696052, 29331828, 264508, 264509, 264510, 264512, 33657402, 60433436, 21906754, 52644296, 67188474, 87168559, 264603, 264661, 264448, 264683, 264288, 264369, 52844229, 264889, 21906765, 21906766, 21906767, 21906768, 21906769, 55811957, 35695917, 265020, 265021, 52844150, 33657023, 264693, 33657182, 35695783, 35696423, 35695855, 52844332, 83373044, 18108367, 87168518, 22279002
2501	84303898 (5001, 5002)	Novel Protein sim. GBank gi14929615 p AA034088.1 AF15183 - (AF15183) CGI-73 protein [Homo sapiens]	Contains protein domain (PF00651) - BTB/POZ domain	dna_mn_bind	65274572, 56182575, 35696286, 22278998, 56994075, 22276997, 60432048, 264259, 29331822, 29331824, 29331828, 29331827, 35696052, 264905, 264908, 264907, 264908, 264909, 58182435, 264510, 264511, 285007, 264910, 264581, 60432228, 33657402, 60433356, 264595, 55812038, 264756, 264598, 87168474, 87168559, 264600, 264601, 264602, 265017, 264604, 265016, 264605, 265019, 18108351, 264448, 264369, 264288, 264766, 18108357, 21906765, 21906766, 21906767, 21906769, 29148628, 35695917, 264692, 33657023, 264629, 35696423, 55811576, 35695855, 264630, 264634, 264635, 264555, 264636, 264638, 264556, 60170394, 83373044, 18108385, 18108387, 87168518, 60432113, 22279002, 264566
2502	80893716 (5003, 5004)	Novel Protein sim. GBank gi13041847 (AC004542) - OXYSTEROL-BINDING PROTEIN-like; similar to P22059 (P1D:8128308) [Homo sapiens]	Contains protein domain (PF01237) - Oxysterol-binding protein	UNCLASSIFIED	65274572, 264907, 56182435, 265007, 264592, 264760, 18108351, 264448, 264369, 264288, 264684, 264686, 55811957, 265021, 264692, 33657109, 263973, 55811576, 264635, 264555, 264556, 264557, 264558, 56182323, 264559, 87168518, 264563, 264462
2503	87678345 (5005, 5006)	Novel Protein sim. GBank gi12186674 emb CAA726391 - (Y11866) BRX protein [Mus musculus]			264905, 264907, 264512, 265008, 265011, 18108351, 264448, 264288, 29148627, 264693, 18108370, 18108374, 18108365

2504	87888708 (5007, 5008)	Novel Protein sim. GBank gll50420[emb]CAA48220 - (X88101) [tg [Rattus norvegicus]			264488, 52644507, 52645156, 52646842, 22278994, 264259, 52645080, 29331822, 29331824, 29331825, 29331826, 29331827, 35686052, 264908, 264908, 52644045, 265009, 60433356, 33657402, 60433438, 284595, 33109954, 87168474, 265017, 265019, 264448, 264288, 264768, 52644229, 21906765, 21906766, 21906767, 21906768, 52644150, 284892, 27488281, 27486262, 27486264, 27486265, 35695763, 35696423, 35695855, 52644332, 56182323, 16108387, 87168518, 60432113, 22278002, 284564
2505	8760559 (5009, 5010)	Novel Protein sim. GBank gll2137562[pil]49635 - mouse		UNCLASSIFIED	264605
2506	91232328 (5011, 5012)	Dhmt protein - mouse		nuclease	264488, 52644507, 52645156, 52646845, 65274572, 22278995, 56994075, 22278996, 22278997, 22278998, 22278999, 264259, 60432049, 29331822, 29331825, 29331826, 29331828, 284509, 56182435, 264112, 284593, 60433358, 55812038, 21906754, 265011, 265017, 265018, 265019, 264605, 264762, 18108351, 264448, 264288, 264768, 21906765, 21906766, 21906767, 21906768, 21906769, 35895917, 265020, 265021, 265022, 60170615, 33657023, 27486264, 18108379, 35695855, 264637, 83373044, 18108385, 87188518, 60432113, 22279000, 264563, 264482, 264565
2507	95318233 (5013, 5014)	Novel Protein sim. GBank gll5174489[ref]NP_006035.1[PKIAA - histone deacetylase 6	Contains protein domain (PF00850) - histone	histone	264488, 263394, 264592, 264595, 264369, 264888, 264768, 35695917, 35896423, 264563
2508	95315505 (5015, 5016)	Novel Protein sim. GBank gll4826433[emb]CAB42889.11 - (AL031447) dJ126A5.2.1 (novel protein) (isoform 1) [Homo sapiens]		UNCLASSIFIED	22278995, 22278999, 60432049, 264259, 29331828, 265006, 265007, 60433438, 33657084, 265010, 285017, 285018, 265019, 18106351, 264448, 18108354, 264368, 18108359, 21906765, 21906769, 55811957, 265020, 265022, 27486261, 33657349, 18106377, 35695855, 60432113, 22279002, 264563, 264565
2509	87613741 (5017, 5018)	Novel Protein sim. GBank gll1263289 (U47856) - fibrin-4 (Araeus diadematus)		UNCLASSIFIED	265007, 265008, 18108357, 264556, 264567

2510	85421378 (5018, 5020)	Novel Protein sim. GBank gll3283537(gb/AAC25762.1) - (AF071059) zinc finger RNA binding protein [Mus musculus]		dna_ma_hind	65274572, 22278994, 22278996, 22278996, 22278996, 60432049, 264259, 29331622, 29331824, 29331825, 66714117, 60432289, 29331826, 33656870, 264908, 66712502, 265007, 264910, 80170631, 60432229, 60433356, 60433438, 21906754, 67166474, 265017, 265018, 264448, 264268, 21906767, 21906798, 21906769, 55611957, 33695917, 265020, 265022, 264691, 33657023, 264693, 65274620, 33657109, 33657182, 27486262, 33657349, 18108370, 35685855, 264555, 58162323, 83373044, 60432113, 22279002, 265017, 21906764, 265020, 264692
2511	87364261 (5021, 5022)	Novel Protein sim. GBank gll4323152(gb/AAD16226.1) - (AF098883) Ets-protein Sp-C [Mus musculus]			
2512	88084771 (5023, 5024)	Novel Protein sim. GBank gll4502075(rejnp_001135.1)pAMFR - autoactive molality factor receptor	Contains protein domain (PF00087) - Zinc finger, C3HC4 type (RING finger)	transport	22278999, 264259, 29331625, 29331826, 29146499, 264907, 264909, 265008, 265008, 264591, 60432229, 21906754, 264783, 264693, 264766, 18108357, 264689, 21906769, 264693, 18108370, 263972, 16108374, 264558, 22279000
2513	95357843 (5025, 5026)	Novel Protein sim. GBank gll3004657 (AF017777) - bobby sex [Drosophila melanogaster]		UNCLASSIFIED	60424178, 52645156, 18106394, 22278994, 35696266, 56994075, 22278996, 29331822, 29331824, 60424269, 29331825, 29331827, 33656970, 80431735, 33857084, 87168559, 265017, 264448, 264369, 58181562, 21906766, 21906767, 21906768, 21906769, 265020, 265021, 33657023, 18108366, 33657109, 27466281, 27466282, 33857349, 18108374, 55610764, 35686423, 58162323, 264558, 16108365
2514	86094578 (5027, 5028)	Novel Protein sim. GBank gll2258437 (AF008197) - syncollin [Rattus norvegicus]		UNCLASSIFIED	264510
2515	87894509 (5029, 5030)	Novel Protein sim. GBank gll375727 (emb)CAA18783 - (AL022272) 648019.7 (olfactory receptor-like protein (rhsm1-3)) [Thomomys talpae]	Contains protein domain (PF00007) - 7 transmembrane receptor (rhodopsin family)	im7	
2516	87766906 (5031, 5032)			UNCLASSIFIED	264259, 29146496, 264905, 264266, 29148629, 35695917, 27466281, 264634, 264091, 29331824, 29331825, 29331826, 29331828, 35696052, 264508, 264509, 264905, 264906, 264907, 264908, 264909, 264511, 264910, 33657402, 264757, 33109954, 265017, 265018, 264605, 264760, 264782, 264783, 264786, 264788, 264789, 33657108, 33657182, 264628, 55611578, 35686423, 264631, 264634, 264637, 264636, 264639, 67166518, 22279002, 264564
2517	87764966 (5033, 5034)	Novel Protein sim. GBank gll4220527 (emb)CAA23000 - (AL035356) putative protein [Arabidopsis thaliana]		UNCLASSIFIED	

2518	94147410 (5035, 5038)	Novel Protein sim. GBank gi4929391 gb AAD34058.1 AF15181 - (AF15181) CGI-61 protein [Homo sapiens]	Contains protein domain (PF00018) - SH3 domain	UNCLASSIFIED	35696286, 264259, 28331822, 29331824, 29331825, 60432289, 29331828, 29331827, 35696052, 29331828, 284807, 264809, 264511, 265007, 60432229, 60433366, 60433438, 55612038, 265010, 265017, 264448, 264288, 264689, 21906766, 21906769, 265022, 52644150, 264693, 18108370, 263972, 284555, 56182323, 83372044, 18108385, 60432113, 264086 284259, 66714177, 29331828, 29331827, 29331828, 264807, 66712502, 265006, 265008, 264594, 265010, 265011, 265018, 264286, 21906769, 265020, 60431528, 55811576, 65274791, 284832, 284555, 264636, 22279002, 264564 284259, 264908, 264910, 284662, 21906769, 265020, 284563
2520	87413235 (5039, 5040)	Novel Protein sim. GBank gi4826722 ref NP_005065.1 pFATP - fatty acid transport protein 4		transport	264486, 264489, 263994, 65274572, 22278998, 22278998, 264259, 29331822, 29331828, 264508, 264905, 264509, 264906, 284907, 66712502, 284511, 265008, 265007, 284581, 284592, 264593, 284594, 284595, 284596, 284681, 284448, 284783, 284682, 264764, 284684, 284369, 284288, 284685, 284686, 21906766, 55811957, 284682, 264693, 27486261, 18108370, 284628, 284629, 18108374, 55811578, 35696423, 35695855, 284632, 284558, 18108385, 65274727, 60432113, 284563, 284564, 284565, 284566, 284567
2521	95316244 (5041, 5042)	Novel Protein sim. GBank gi15174489 ref NP_006035.1 pKIAA - histone deacetylase 6	Contains protein domain (PF00050) - histone histone deacetylase family	histone	264486, 264489, 263994, 65274572, 22278998, 22278998, 264259, 29331822, 29331828, 264508, 264905, 264509, 264906, 284907, 66712502, 284511, 265008, 265007, 284581, 284592, 264593, 284594, 284595, 284596, 284681, 284448, 284783, 284682, 264764, 284684, 284369, 284288, 284685, 284686, 21906766, 55811957, 284682, 264693, 27486261, 18108370, 284628, 284629, 18108374, 55811578, 35696423, 35695855, 284632, 284558, 18108385, 65274727, 60432113, 284563, 284564, 284565, 284566, 284567
2522	87754052 (5043, 5044)	Novel Protein sim. GBank gi4580011 gb AAD24201.1 U61002 - (U61002) TRAF4 associated factor 1 [Homo sapiens]		transport	264486, 22278997, 20281171, 21906754, 35695817, 263967, 263976, 263981, 20281169
2523	95340467 (5045, 5046)	Novel Protein sim. GBank gi1809327 (U76374) - skm-	Contains protein domain (PF01753) - MTND finger		56694075, 22278996, 35696286, 22278997, 22278998, 22278999, 284259, 29331822, 29331824, 29331826, 29331828, 29331830, 56182435, 264512, 265008, 60170831, 33657402, 265010, 87168559, 285019, 284288, 21906765, 21906769, 35695917, 265020, 265021, 265022, 52644150, 284691, 33657023, 33657109, 27486261, 35696423, 65274791, 284559, 63373044, 56526466, 87168518, 264567
2524	95340469 (5047, 5048)	Novel Protein sim. GBank gi1809327 (U76374) - skm-			

2525	84126828 (5049, 5050)	Novel Protein sim. GBank gij2073564 (U80223) - eukaryotic initiation factor eIF-2, alpha kinase. DGCNZ [Drosophila melanogaster]		kinase	264488, 22278997, 22278999, 60432048, 80432289, 29331828, 264805, 265008, 55812038, 21906754, 265019, 264369, 21906785, 21906766, 21906767, 21906769, 35695917, 265020, 265021, 33857109, 60431528, 83373044, 80432113, 22279000, 22279002, 264585
2528	95289404 (5051, 5052)	Novel Protein sim. GBank gik458828jdbjBAA78338.1 - (AB023209) KIAA0892, protein [Homo sapiens]	Contains protein domain (PF00238) - ribosomal protein L14	ribosomal prot	80424179, 264768, 264687, 264769, 264689, 65274572, 21906767, 58182575, 21906786, 21906769, 55811957, 22278994, 22278995, 35695917, 22278998, 22278997, 265020, 22278998, 265021, 22278999, 265022, 264680, 264891, 60432048, 264259, 264097, 33657023, 29331822, 29331824, 80432289, 29331826, 29331827, 29331828, 27486282, 264508, 264509, 264805, 264807, 18108370, 86712502, 80431528, 264828, 264809, 18108372, 18108374, 56182435, 18108378, 55810784, 55811578, 35696423, 35695855, 285006, 265007, 284512, 285008, 265009, 264634, 264835, 60431850, 264638, 264555, 264592, 80431735, 284638, 33657402, 56182323, 60433358, 80433438, 264585, 55812038, 264598, 264758, 83373044, 52846317, 18108385, 33657084, 18108387, 55811388, 85274727, 58526488, 87168518, 80432113, 265017, 22279000, 285018, 285019, 264584, 18108351, 284448, 264586, 264288, 264488, 264587, 264766, 264510, 264512, 264630, 264591, 264592, 284259, 264594, 264595, 264603, 264805, 18108351, 264585, 264389, 18108354
2527	88094580 (5053, 5054)	Novel Protein sim. GBank gij2258437 (AF008197) - syncollin [Rattus norvegicus]		UNCLASSIFIED	55812038, 285017, 264889, 35695917, 35695763, 60431528, 60432113, 22279002
2528	88076380 (5055, 5056)	Novel Protein sim. GBank gij2085788 (AC002086) - similar to zinc finger 5 protein from Gallus gallus, U51640 (PIDg1393185) [Homo sapiens]	Contains protein domain (PF00651) - BTB/POZ domain	dna_ga_bind	264808, 264768, 265020, 265021, 18108383
2529	88870828 (5057, 5058)	Novel Protein sim. GBank gik3786433 (AF098505) - similar to Arabidopsis thaliana male sterility protein 2 (SW:Q08891) [Caenorhabditis elegans]		synthase	264369, 264556
2530	80259978 (5058, 5060)			UNCLASSIFIED	28331822, 28331824, 60432289, 264508, 264509, 284808, 265011, 264769, 21906768, 33857023, 87168518, 22279000
2531	87768831 (5081, 5082)			UNCLASSIFIED	264593
2532	87419778 (5083, 5084)	Novel Protein sim. GBank gij2864625jembCAA168721 - (AL021811) putative protein [Arabidopsis thaliana]		UNCLASSIFIED	284555
2533	87000255 (5085, 5086)	Novel Protein sim. GBank gik437181 (U02289) - GTPase-activating protein [Caenorhabditis elegans]		UNCLASSIFIED	

2534	87332322 (5087, 5088)	Novel Protein sim. GBank gll3452473 (AF084205) - serine/threonine protein kinase TAO1 [Rattus norvegicus]	UNCLASSIFIED	264259, 35696052, 264905, 265017, 21908769, 265020, 265022, 33657109, 22279000
2535	91225056 (5088, 5070)	Novel Protein sim. GBank gll448311 (emb) (CAB37992) - (AL031432) dJ485N24.2.1 (PUTATIVE novel protein) (soform 1) [Homo sapiens]		85274572, 356968286, 604322289, 29331828, 66712502, 265006, 60432229, 285017, 265018, 285019, 264288, 264369, 264689, 21906768, 265020, 265021, 264636, 60170394, 22279002
2536	94218540 (5071, 5072)	Novel Protein sim. GBank gll728838 (p39193) ALU8 - III ALU SUBFAMILY SP WARNING ENTRY IIII	kinase	18108398, 56182375, 35696286, 22278997, 22278999, 60432049, 264259, 29331824, 29331828, 29331827, 29331828, 264905, 264511, 285009, 264910, 264596, 52646317, 18108351, 264681, 264883, 18108354, 264288, 264687, 264789, 264689, 21906765, 21906766, 21906787, 265021, 52645129, 33657109, 18108374, 18108380, 58182323, 18108381, 18108388, 87168518, 60432113, 22279000, 22279002, 264567, 18108391
2537	95422283 (5073, 5074)	Novel Protein sim. GBank gll4557028 (re) NP_000913.1 (pHERC - guanine nucleotide exchange factor p532	ubiquitin	65274572, 35696286, 29331822, 29331825, 29331827, 29331828, 35696052, 264906, 66712502, 264909, 265008, 285011, 264760, 264288, 264885, 35695917, 60170615, 264681, 33857023, 65274620, 33857109, 18108374, 35696423, 35695855, 264638, 264558, 60170394, 56182323, 83373044
2538	36853454 (5075, 5076)		UNCLASSIFIED	22278998, 22278999, 29331822, 29331825, 29331828, 29146499, 264908, 264112, 60170831, 87188559, 264804, 265019, 264885, 264788, 87168518, 22279000, 264565, 264568
2539	84144818 (5077, 5078)		UNCLASSIFIED	22278997, 29331828, 285008, 265009, 264758, 265010, 18108351, 264683, 264288, 21906765, 35695917, 265020, 18108374, 264567
2540	94218545 (5078, 5080)	Novel Protein sim. GBank gll1382647 (p) [S53876 - sex-regulated protein Janus A - fruit fly (Drosophila pseudobscura)]	UNCLASSIFIED	264488, 264768, 264889, 264511, 20281171, 264834, 264835, 264681, 264639, 29331824, 264603, 264804, 264905, 264907, 264908, 264768
2541	85308238 (5081, 5082)	Novel Protein sim. GBank gll1711858 (p) [P54797] TIO MOUSE - SER/TMR-RICH PROTEIN TIO IN DCCR REGION	UNCLASSIFIED	

2542	86298162 (5083, 5084)	Novel Protein sim. GBank gi 5225320 gb ACD0850.1 AF06310 - (AF083107) sirutin type 2 [Homo sapiens]	Contains protein domain (PF00220) - Neurohypophyseal hormones, N- terminal Domain	UNCLASSIFIED	264488, 16108394, 52846365, 52846842, 65274572, 22276994, 35696266, 22276996, 264259, 52845080, 29331822, 29331624, 29331827, 35696052, 33856970, 284907, 284909, 52844045, 264510, 265006, 284512, 265007, 265008, 265009, 264910, 60431735, 52846317, 52844286, 265010, 265011, 265018, 265019, 18108351, 264683, 264286, 264685, 284687, 52844229, 284769, 21906766, 21906767, 21905769, 52844150, 33657023, 33657109, 52645128, 33657182, 27486261, 27486264, 33657348, 3569583, 18108374, 35696423, 35695855, 264631, 264634, 264635, 284558, 63373044, 18108385, 16108387, 87168516, 264563, 264564
2543	84139086 (5085, 5086)	Novel Protein sim. GBank gi 5419857 emb CA946374.1 - (AL096723) hypothetical protein [Homo sapiens]	Contains protein domain (PF00076) - RNA recognition motif (a.k.a. RRM, RBD, or RNP domain)		65274572, 56162575, 22276999, 264259, 29331826, 264907, 264510, 264511, 264592, 264595, 264764, 264369, 264288, 264884, 264766, 264689, 21906765, 21906767, 21906789, 60170615, 264692, 264693, 55811576, 65274791, 264636, 264556, 18108381, 60170394, 264639, 18108385, 60432113, 22279000
2544	84218549 (5087, 5088)	Novel Protein sim. GBank gi 2498110 sp Q63191 AEGP_RAT - APICAL ENDOSOMAL GLYCOPROTEIN PRECURSOR	Contains protein domain (PF00629) - MAM domain.	glycoprotein	18108397, 52846365, 22278997, 264259, 60432048, 29331822, 29331825, 29331826, 29331827, 29331828, 264905, 264906, 265006, 265007, 265008, 67168559, 285017, 265018, 265019, 16108351, 264446, 264688, 264687, 264689, 21906765, 265020, 285021, 18108370, 18108374, 18108376, 18108381, 18108385, 18108387, 56526486, 22279000, 264482, 264583, 264587
2545	67742845 (5089, 5090)	Novel Protein sim. GBank gi 3327046 db BA315911 - (AB014516) KIAA0618 protein [Homo sapiens]			29331825, 284906, 285009, 60170831, 265017, 264369, 21906767, 60170615, 264692, 33657109
2546	68083861 (5091, 5092)	Novel Protein sim. GBank gi 2996032 (AF054386) - brain finger protein [Rattus norvegicus]	Contains protein domain (PF00097) - Zinc finger, C3HC4 type (RING finger)	UNCLASSIFIED	29331824, 265007, 22278002

2547	94143869 (5083, 5094)	Novel Protein sim. GBank g14928607 b b AAD34064_1 AF15162 - (AF151627) CGI-69 protein [Homo sapiens]	Contains protein domain (PF00153) - transport	264486, 18108394, 52646842, 18108397, 56182575, 22278995, 56864075, 22278998, 22278997, 22278999, 264259, 28331822, 29331824, 29331826, 60432289, 29331827, 35696052, 28331826, 284104, 264508, 264905, 264906, 264908, 66712502, 264909, 56182435, 265004, 265007, 264512, 265006, 265009, 60170631, 60432229, 60431735, 264594, 60433438, 21908754, 52646317, 265010, 265011, 264600, 264601, 265018, 265019, 264760, 18108351, 264682, 264446, 264288, 264368, 264684, 264686, 264687, 56181582, 264688, 264689, 21908785, 21908786, 21908787, 21908788, 29146627, 21908789, 55611857, 265020, 265021, 265022, 264690, 264691, 16106362, 264692, 264693, 27466261, 18108370, 18108374, 55810764, 55611576, 35686423, 35686455, 264635, 264636, 264555, 264637, 263981, 264557, 18108380, 264638, 56182323, 264558, 264559, 63373044, 18108385, 67168516, 22279002, 264564, 264566, 264486
2548	88179078 (5085, 5086)		UNCLASSIFIED	264488, 18108394, 52646845, 22278994, 35686286, 56894075, 22278997, 22278998, 264259, 29331822, 29331824, 29331825, 29331826, 60432289, 29331827, 29331828, 56182435, 264511, 265007, 264512, 60433356, 67168559, 264684, 264369, 52644229, 265021, 33657023, 264692, 18108374, 52644332, 264557, 18108380, 18108381, 18108382, 18108384, 18108385, 60432113, 22279000, 22279002, 264563, 264567
2549	94198683 (5097, 5098)	Novel Protein sim. GBank g1728837 sp P39194 ALU7_HUMAN - III ALU SUBFAMILY SQ WARNING ENTRY III	Contains protein domain (PF00412) - struct	56182575, 22278998, 22278997, 22278998, 22278999, 264259, 264506, 264606, 29331830, 265009, 265010, 265018, 264688, 21908784, 21908785, 21908786, 21908787, 21908788, 265020, 265021, 52644150, 264691, 18108386, 60431802, 18108378, 35696423, 56182323, 18108387, 264567 56182575, 29331822, 264105, 264512, 18108351, 35695917, 264637, 264638
2550	87778564 (5089, 5100)	Novel Protein sim. GBank g12143888 p I52523 - nucleoporin p62 homolog - rat (fragment)	UNCLASSIFIED	

2551	95308400 (5101, 5102)	Novel Protein sim. GBank g14337103 gb AAD180791 - (Ae129756) NC28 (Homo sapiens)	Contains protein domain (PF00036.1) - alpha/beta hydrolase fold	UNCLASSIFIED	18108398, 85274517, 22278995, 22278998, 22278999, 264259, 28331822, 28331824, 28331825, 60432289, 28331828, 28331827, 264905, 56182435, 265007, 50433438, 55812038, 21906754, 85274444, 265017, 265016, 264605, 285019, 264288, 21908766, 21908768, 21906769, 285020, 60170815, 264693, 33657109, 35698423, 284638, 58182323, 83373044, 22279000
2552	95332820 (5103, 5104)			UNCLASSIFIED	58182575, 35696288, 28331824, 28331826, 35696052, 28331828, 284508, 284907, 56182435, 285008, 264591, 33109954, 284760, 55811957, 35895917, 33857023, 33857109, 18108374, 55811576, 35698423, 35695855, 58182323, 264558
2553	95308243 (5105, 5106)	Novel Protein sim. GBank g11711858 sp P4787 T10_MOUSE - SER/THR-RICH PROTEIN T10 IN DSCR REGION		UNCLASSIFIED	264686, 264488, 283978, 264788, 29331826, 35696052, 35698423, 284601, 264511, 284602, 264910, 264634, 264760, 264555, 284782, 264908, 284592, 284891, 284568, 264808, 264684, 284587, 264909, 264768
2554	87761520 (5107, 5108)	Novel Protein sim. GBank g1728835 sp P38192 ALU5_HUMAN - IIII ALU SUBFAMILY SC WARNING ENTRY IIII		cadherin	22278897, 29331822, 264508, 21908769, 33657023, 33657109, 58182323
2555	87627551 (5109, 5110)	Novel Protein sim. GBank g1484319 emb CAB3260.1 - (AL050084) hypothetical protein [Homo sapiens]		nuclease	29331824, 263872
2556	87645533 (5111, 5112)	Novel Protein sim. GBank g14105984 (AC003038) - R30923_1 [Homo sapiens]	Contains protein domain (PF00514) - Armadillo/beta-catenin-like repeats	UNCLASSIFIED	22278998, 264509, 33657402, 284683, 264884, 264768, 264689, 33657023, 33657109, 35695855, 264558, 264567
2557	78437803 (5113, 5114)				264595
2558	87617591 (5115, 5116)	Novel Protein sim. GBank g1119110 sp P03211 EBN1_EBV - EBNA-1 NUCLEAR PROTEIN		UNCLASSIFIED	22278897, 28331824, 66714117, 28331825, 284906, 264511, 265018, 264448
2559	88086382 (5117, 5118)	Novel Protein sim. GBank g1453899 emb CAB39519.1 - (AL049481) ALG1-like protein [Arabidopsis thaliana]		UNCLASSIFIED	22278997, 28331822, 28331826, 60433356, 285011, 264288, 264765, 264766, 264769, 21908765, 21908768, 60432113, 264482
2560	87784530 (5119, 5120)	Novel Protein sim. GBank g15051389 emb CAB44995.1 - (AL078630) 573K1.3 (nm17M1-4) (novel 7 transmembrane receptor (rhodopsin family) (olfactory receptor like protein)) [Mus musculus]	Contains protein domain (PF00001) - 7 transmembrane receptor (rhodopsin family)		
2561	88176575 (5121, 5122)	Novel Protein sim. GBank g15326825 gb AAD42056.1 AF044953 (AFO44953) NADH:ubiquinone oxidoreductase PGIV subunit [Homo sapiens]		UNCLASSIFIED	22278895, 35696286, 22278996, 22278997, 22278998, 22278999, 264259, 28331822, 28331824, 28331828, 29331827, 29331828, 285007, 80432229, 87168559, 285017, 265018, 285019, 264689, 21908768, 21908769, 35695917, 285020, 33657023, 33657109, 18108374, 264634, 284559, 18108385, 87188518, 22279002

2582	87645539 (5123, 5124)	Novel Protein sim. GBank g14106984 (AC003038) - R30823_1 [Homo sapiens]		UNCLASSIFIED	568964075, 222769896, 222769897, 222769898, 264259, 29331822, 60432289, 33657402, 60433358, 21906765, 55811957, 60170815, 33657023, 264693, 35695855, 87168518
2583	88095487 (5125, 5126)	Novel Protein sim. GBank g14886447 (emb CAB43371.1 - (AL050270) hypothetical protein [Homo sapiens]		UNCLASSIFIED	264488, 35696286, 22276989, 264259, 29331822, 29331824, 35698052, 264508, 264807, 264908, 264909, 52644045, 264510, 264511, 265009, 264810, 264591, 284593, 33657402, 265017, 265018, 265019, 16108351, 264686, 21906767, 21906768, 55811957, 35695917, 265020, 264691, 264693, 27466262, 264628, 18108374, 35696423, 35695855, 264632, 264634, 264835, 264639, 264558, 18108384, 87168518, 22279000, 22279002, 264482, 264563, 264565, 264566, 264486
2584	80502763 (5127, 5128)	Novel Protein sim. GBank g11352944 (sp P47179 YJP_YEAST - HYPOTHETICAL 118.4 KD PROTEIN IN BAT2-DAL5 INTERGENIC REGION - PRECURSOR		sulfotransferase	29331822, 265007, 265010, 265019, 264769, 55811576, 56182323
2585	85530906 (5129, 5130)	Novel Protein sim. GBank g1628012 (pf A53933 - myosin I myr 4 - rat	Contains protein domain (PF00063) - Myosin head (motor domain)	UNCLASSIFIED	66714117, 264909, 263978, 264632, 18108370, 35695855, 264558, 264558, 18108383
2586	80224856 (5131, 5132)	Novel Protein sim. GBank g1628012 (pf A53933 - myosin I myr 4 - rat	Contains protein domain (PF00063) - Myosin head (motor domain)	struct	265020, 60170815
2587	88143590 (5133, 5134)	Novel Protein sim. GBank g1468009 (sp P34540 YNU4_CAEEL - HYPOTHETICAL 33.8 KD PROTEIN R10E11.4 IN CHROMOSOME III		synthase	60424178, 18108384, 58181888, 56994075, 22276989, 264480, 264259, 29331822, 56182181, 29331824, 60424289, 29331825, 29331826, 29331828, 264509, 29331830, 265007, 265008, 265009, 33657402, 265010, 265011, 265018, 264448, 264683, 264288, 18108354, 264769, 21906766, 21906767, 35695917, 265021, 33657023, 16108362, 33657109, 33657182, 35695763, 60431528, 55810764, 18108379, 83373044, 18108385, 60432113, 264482

2569	95313764 (5137, 5138)	Novel Protein sim. GBank gij2599560jgijAB84166.11 - (AF029874) basic leucine zipper protein LZP [Homo sapiens]	Contains protein domain (PF00170) - bZIP transcription factor	transcription factor	18108394, 56182575, 56181686, 22278995, 22278996, 56994075, 35696286, 22278997, 22278998, 22278999, 264259, 60432049, 29331822, 56182181, 29331824, 29331825, 60432289, 29331828, 29331827, 29331828, 35696052, 264905, 58182435, 265008, 265009, 264910, 60432229, 264592, 60433358, 60433438, 21908754, 87188559, 265017, 265018, 265019, 264682, 264448, 264288, 21908765, 21908766, 21908767, 21908768, 29146827, 21908769, 35695917, 265021, 265022, 52644150, 264690, 264691, 264692, 264693, 65274620, 263967, 35695763, 20281058, 263974, 18108374, 55810764, 35695423, 35695855, 264558, 18108381, 56182323, 18108382, 83373044, 18108385, 56526488, 87168518, 22278000, 22278002, 264568
2570	94136754 (5138, 5140)	Novel Protein sim. GBank gij4758954jgijFP_004567.1jgijPP2 - protein phosphatase 2 (formerly 2A), regulatory subunit B (PR 52), beta isoform	Contains protein domain (PF01240) - Protein phosphatase 2A regulatory subunit PR55	phosphatase	22278996, 29331822, 29331824, 66714117, 29331825, 60432289, 29331827, 35696052, 264907, 264510, 265007, 265009, 264758, 33109954, 285019, 264688, 264689, 265020, 265021, 264691, 264692, 264693, 18108374, 35695855, 264634, 56182323, 264639, 60170394, 83373044, 22278002, 264482, 264508, 264905, 264907, 264628, 16108351, 264555, 264556, 264557, 264558, 264559
2571	87733750 (5141, 5142)	Novel Protein sim. GBank gij732218jgijPF34609jY060 CAEL - HYPOTHETICAL 128.6 KD PROTEIN ZK1098.10 IN CHROMOSOME III		struct	
2572	87927560 (5143, 5144)	Novel Protein sim. GBank gij4684319jemb[CAB43260.11 - (AL050084) hypothetical protein [Homo sapiens]		UNCLASSIFIED	35696286, 29331827, 35696052, 264100, 264104, 264110, 264592, 21908754, 29148627, 29148628, 263972, 263974, 16108374, 263976, 35695655, 60170394, 264559, 18108385

2573	85313828 (5145, 5146)	Novel Protein sim. GBank gI389138pPoz745IC10A, HUMAN - COMPLEMENT C10 SUBCOMPONENT, A CHAIN PRECURSOR	Contains protein domain (PF00386) - C1q domain	- complement	264488, 60424178, 65274572, 56182575, 56181686, 22278893, 56894075, 22278897, 60432049, 264259, 29331822, 29331824, 29331825, 60432288, 29331826, 29331827, 29331828, 264104, 264107, 264508, 264806, 29331830, 264909, 264510, 265006, 264512, 265008, 265009, 264910, 264581, 264592, 60432229, 264593, 60433356, 264594, 60433438, 264595, 56812038, 264759, 21906754, 87168474, 265010, 265011, 87168559, 265017, 265018, 265019, 264781, 264782, 264783, 264784, 264389, 264288, 264685, 264786, 264888, 264687, 264686, 264769, 56181562, 264689, 21906765, 21906766, 21906767, 29148627, 21906768, 21908789, 265020, 265021, 265022, 60170615, 264690, 52844150, 264691, 264692, 33657023, 65274620, 18108365, 18108368, 27486265, 60431602, 264628, 60431528, 263978, 65274791, 35695855, 20281071, 60431850, 264637, 264638, 264558, 264639, 56182323, 60170394, 83373044, 18108384, 87188518, 80432113, 264482, 264564, 264585, 264586, 264587, 22278895, 264258, 60432289, 29331827, 29331828, 33658970, 264908, 265008, 264910, 264581, 33657402, 265018, 265019, 264448, 264764, 264369, 264288, 18108357, 21906765, 21906766, 21906768, 55811857, 60170615, 264691, 33657023, 264693, 33657109, 33657182, 27488281, 27488284, 33657349, 264636, 264555, 83373044, 18108365, 264482
2574	84748814 (5147, 5148)	Novel Protein sim. GBank gI3334892 (AC005306) - R27218_1 (Homo sapiens)	Contains protein domain (PF00651) - BTB/POZ domain	UNCLASSIFIED	264482, 264564, 264585, 264586, 264587, 22278895, 264258, 60432289, 29331827, 29331828, 33658970, 264908, 265008, 264910, 264581, 33657402, 265018, 265019, 264448, 264764, 264369, 264288, 18108357, 21906765, 21906766, 21906768, 55811857, 60170615, 264691, 33657023, 264693, 33657109, 33657182, 27488281, 27488284, 33657349, 264636, 264555, 83373044, 18108365, 264482
2575	87754408 (5149, 5150)	Novel Protein sim. GBank gI4929729gIb)AA034125, IAF15188 - (AF15188) CGI-130 protein (Homo sapiens)		UNCLASSIFIED	264810, 264601, 264663, 264689, 264080
2576	95357861 (5151, 5152)	Novel Protein sim. GBank gI4680881)gIb)AAD27730, IAF13295 - (AF13295) CGI-21 protein (Homo sapiens)	Contains protein domain (PF00442) - Ubiquitin carboxyl-terminal hydrolases family 2	ubiquitin	264259, 35696052, 264806, 60433438, 264881, 18108351, 264288, 52844150, 264628, 35696423
2577	86898621 (5153, 5154)	Novel Protein sim. GBank gI4337103)gIb)AAD180781 - (AF128756) NG28 (Homo sapiens)		UNCLASSIFIED	29331825, 265016, 265019, 264665
2578	87786941 (5155, 5156)			UNCLASSIFIED	264488, 264806, 264808, 264910, 264586, 264803, 264604, 264605, 264768, 21906769, 264828, 264830, 264834, 264639, 264563
2579	87292878 (5157, 5158)			UNCLASSIFIED	29331822, 29331824, 264767

2580	86188768 (5159, 5160)	Novel Protein sim. GBank gij2568628 (AC003080) - Similar to KIAA0298: 80% similarity to AB002297 (P1D:g2224339) [Homo sapiens]			265007, 265018, 264762
2581	87899048 (5181, 5182)	Novel Protein sim. GBank gij1406642gpiAACD20049) - (AF131609) Unknown [Homo sapiens]	Contains protein domain (PF00595) - PDZ domain (Also known as DHR or GGF).	collagen	56894075, 29331824, 29331826, 29331828, 264805, 60433358, 60433438, 264758, 87168559, 21906769, 265022, 35685855, 263981
2582	87788789 (5183, 5184)	Novel Protein sim. GBank gij2738397 (AC002505) - putative phosphatidylinositol-4-phosphate 5-kinase [Arabidopsis thaliana]		eph	264488, 264907, 264908, 264910, 264764, 264684, 264766, 264836, 264555, 264565
2583	91220950 (5165, 5166)	Novel Protein sim. GBank gij437812gembCAA18521.1) - (AL021578) d4453C12.2 (similar to transcription factor RBP-L) [Homo sapiens]	Contains protein domain (PF00047) - Immunoglobulin domain	transcription factor	56181886, 264259, 264510, 264512, 264591, 264592, 264593, 264594, 264595, 264596, 264603, 264629, 55810764, 264630, 264637, 264565
2584	60430941 (5167, 5168)			UNCLASSIFIED	264908, 264910, 264786, 264693, 18108374, 55811576, 58182323
2585	80436128 (5189, 5170)	Novel Protein sim. GBank gij2736151 (AF021935) - myosin dystrophy kinase-related Cdc42-binding kinase [Rattus norvegicus]		kinase	264768
2586	91226136 (5171, 5172)				22278998, 264259, 29331822, 29331824, 29331827, 29331828, 264906, 265007, 265009, 264591, 60433358, 33667402, 265016, 264762, 264288, 21908766, 21908787, 21908769, 265022, 264891, 63373044, 58528486, 22278002
2587	80430943 (5173, 5174)				264808, 265019, 264768, 264693, 55811576, 58182323
2588	80074385 (5175, 5176)			UNCLASSIFIED	264564
2589	85516607 (5177, 5178)	Novel Protein sim. GBank gij021590gembCAA714151 - (Y10389) nuclear protein [Xenopus laevis]		UNCLASSIFIED	35696052, 264905, 264908, 264907, 264908, 264909, 265009, 265018, 264789, 35696423, 264636
2590	87054526 (5179, 5180)	Novel Protein sim. GBank gij2104889 (U82783) - alpha glucosidase II, alpha subunit [Mus musculus]	Contains protein domain (PF01055) - Glycosyl hydrolases family 31	glucosylase	22278995, 29331830, 265008, 265010, 265017, 264639
2591	84192167 (5181, 5182)	Novel Protein sim. GBank gij5702202gpiAACD47199.1)AF12918 - (AF129186) long-chain acyl-CoA synthetase 5 [Homo sapiens]		eph	264259, 29331822, 264106, 264906, 56182435, 265007, 265008, 33108954, 264448, 55811957, 265020, 18108370, 55811576, 22279002

2592	95332648 (5183, 5194)	Novel Protein sim. GBank gi13024998 sp Q60806 YAB1_MOUSE - HYPOTHETICAL HEART PROTEIN		transport	18108397, 56182575, 35686286, 56984075, 284259, 29331822, 29331824, 29331826, 60432289, 29331827, 29331828, 284906, 284908, 285007, 285008, 284910, 80432228, 284594, 80433358, 80433438, 55612036, 18108348, 21906754, 265011, 87166558, 285017, 265019, 284764, 284369, 284286, 284768, 265021, 60170615, 33657023, 33657109, 264628, 35686423, 35685855, 284557, 284638, 80170394, 58182323, 63373044, 56526486, 87188518, 284563, 284482, 284565
2593	87754416 (5185, 5186)	Novel Protein sim. GBank gi14929729 p A0D34125.1 AF15188 - (AF15188) CGI-130 protein [Homo sapiens]		tm7	22278999, 29331825, 284756, 21906754, 52646317, 265010, 18108351, 284288, 264369, 21906758, 264693, 18108370, 284637, 284638, 284482
2594	95305758 (5187, 5188)	Novel Protein sim. GBank gi14929587 p A0D34054.1 AF15181 - (AF15181) CGI-59 protein [Homo sapiens]		UNCLASSIFIED	264488, 18108398, 56182575, 35686286, 22278997, 264093, 264259, 29331822, 29331825, 66714117, 29331828, 284905, 264909, 52644045, 56182435, 264510, 264512, 265007, 264757, 21906754, 87168474, 265017, 264760, 264448, 264764, 284288, 284768, 284689, 21906788, 33657109, 263875, 263877, 264634, 264558, 60170394, 56182323, 56526486, 264462, 264563, 264564, 264568, 284587
2595	79561676 (5189, 5190)			UNCLASSIFIED	264892
2596	87538637 (5191, 5192)	Novel Protein sim. GBank gi14309681 p A0D15478 - (ACD06830) R33423.1 [Homo sapiens]		UNCLASSIFIED	22278998, 264259, 265018, 284446, 285021, 80431528
2597	94784089 (5183, 5194)			UNCLASSIFIED	284905, 284509, 284908, 284762, 284768, 35695917, 35695855, 284635, 284636, 83373044, 264488
2598	88094948 (5185, 5189)	Novel Protein sim. GBank gi11001351 p BAA10838 - (CB4008) hypothetical protein [Synchocystis sp.]		UNCLASSIFIED	22278998, 284259, 29331824, 87168474, 284683, 21906768, 35695917, 284691, 33657023, 33657109, 18108370, 18108374, 264564, 284565
2599	87842889 (5197, 5198)	Novel Protein sim. GBank gi13941737 (AF108719) - BA12 [Mus musculus]		MHC	264768, 264769, 21906768, 33657182, 35695763, 18108370, 18108374, 284635, 284636, 56526486, 22278900, 264586
2600	87787848 (5189, 5200)	Novel Protein sim. GBank gi14263521 p A0D15347 - (ACD04044) putative WD-repeat protein [Arabidopsis thaliana]	Contains protein domain (PF00400) - WD domain, G-beta repeat	kinasereceptor	35698286, 264093, 264288, 21906759, 35698423, 35695855

2801	91243070 (5201, 5202)	Novel Protein sim. GBank g1728637isp194JALUT_HUMAN - III ALU SUBFAMILY SQ WARNING ENTRY III	Kinase	56182575, 22276999, 264259, 29331822, 29331824, 60432288, 29331826, 29331827, 264908, 265007, 285008, 264591, 60433356, 33657402, 60433438, 21906754, 265011, 265018, 295019, 16108351, 264446, 264369, 21906769, 265020, 60170615, 264693, 33657109, 18108370, 18108376, 56182323, 18108381, 18108385, 22276902, 264563 60433438, 21906754, 87165559, 264601, 264389, 264288, 21906767
2802	88180022 (6203, 6204)	Novel Protein sim. GBank g14406632gpbA020047 - (AF131801) Unknown (Promo sapiens)		264468, 65274572, 22276995, 22276996, 56994075, 22276997, 22276998, 22276999, 6043049, 264259, 29331822, 29331824, 60432288, 29331826, 29331827, 29331828, 35696052, 58182435, 264113, 265008, 265009, 60433356, 264757, 60433436, 264759, 33657084, 87168474, 265010, 265011, 87168559, 265017, 265018, 265019, 264448, 264683, 16108354, 264288, 264767, 264689, 21906765, 21906766, 21906767, 21906768, 21906769, 55811957, 285020, 265021, 265022, 60170615, 264691, 33657023, 264693, 33657109, 27486282, 18108374, 35696423, 65274791, 35696555, 264555, 264636, 264637, 58182323, 83373044, 56526486, 87168516, 60432113, 22279000
2804	94678601 (5207, 5208)	Novel Protein sim. GBank g1545403qre1NP_006489.1pRRP2 - RAS-related on chromosome 22	oncogene	264259, 35696052, 264508, 264908, 264907, 264608, 264609, 264510, 264512, 265008, 264910, 33657402, 264804, 264805, 264762, 264763, 264682, 264764, 264683, 264768, 264769, 264689, 33657023, 264693, 16108365, 264828, 35696423, 264631, 264832, 264834, 264635, 264637, 16108381, 264839, 63373044, 264965
2805	94318756 (5208, 5210)	Novel Protein sim. GBank g13628745pdlBAA333661 - (AB013721) mltisugumh 23 (Oryzodagus carinatus)	UNCLASSIFIED	22276998, 264490, 60432049, 264259, 60432288, 264909, 265008, 60433356, 60433438, 264756, 21906754, 265010, 265011, 265018, 264681, 16108351, 264288, 264768, 264685, 21906765, 21906766, 21906768, 21906769, 264691, 264692, 264693, 65274791, 264634, 264555, 264636 22276998, 264510, 264512, 265009, 264766, 22279002, 264566
2808	67746406 (5211, 5212)			

2807	67627742 (5213, 5214)	Novel Protein sim. GBank g14826626 gb AAD30202.1 - (AF135022) mediator [Homo sapiens]			29331822, 29331625, 29331626, 29331827, 29331628, 264906, 264907, 264908, 68712502, 264826, 56162435, 55812038, 265010, 265017, 265016, 265018, 264788, 264689, 21906765, 55811957, 265020, 265022, 264692, 33657023, 264693, 33657109, 18108370, 264639, 56162323
2808	81734786 (5215, 5216)	Novel Protein sim. GBank g12226005 (U9973) - ORF2: function unknown [Homo sapiens]			264509, 264907, 264908, 264592, 264758, 264631
2809	94843791 (5217, 5218)	Novel Protein sim. GBank g13024889 P56524 Y288_HUMAN - HYPOTHETICAL PROTEIN KIAA0288 (HAG116)	Contains protein domain (PF00650) - histone Histone deacetylase family	transcript factor	264486, 65274572, 35696268, 22276997, 22278999, 60432049, 264259, 58162181, 29331624, 29331625, 29331628, 80432289, 29331627, 29331628, 264905, 264907, 80433356, 60433438, 55812038, 265011, 87168559, 265017, 265018, 264446, 264765, 264286, 264766, 264689, 21906785, 21906787, 21906789, 285020, 285021, 264691, 264692, 33657109, 27488281, 18108370, 65274791, 264636, 264556, 56182323, 18108385, 56526486
2810	86177654 (5219, 5220)	Novel Protein sim. GBank g14336655 gb AAD179891 - (AF106473) leucine-rich-domain inter-acting protein 1; LEA1 inter-acting protein 1; LEAP1 [Mus musculus]		transcript factor	18108394, 22278994, 56994075, 60432049, 284258, 29331822, 29331825, 60432289, 29331827, 264107, 264109, 264805, 56182435, 264112, 265008, 265007, 265008, 265009, 80433356, 60433438, 265011, 87188559, 265017, 264448, 264682, 264764, 264286, 265021, 33657023, 283987, 33657182, 27488281, 18108374, 263978, 55811578, 264638, 87168518, 60432113
2811	87428880 (5221, 5222)	Novel Protein sim. GBank g13876761 emb CAA92984 - (Z68760) predicted using GeneFinder: Similarity to Mouse FK506-binding protein (SW:FKB3_MOUSE) [Caenorhabditis elegans]	Contains protein domain (PF00254) - isomerase FKB-type peptidyl-prolyl cis-trans isomerases	isomerase	22278899, 265017, 264684, 21906768, 22279000
2812	67771198 (5223, 5224)	Novel Protein sim. GBank g15679138 gb AAD46874.1 AF16093 - (AF180934) BCDNA.LD14189 [Drosophila melanogaster]		transport	265009, 264910, 264759, 265017, 21906767, 16106365, 16106386, 60432113
2813	79481496 (5225, 5226)	Novel Protein sim. GBank g15533081 gb AAD45003.1 AF18116 - (AF161181) P55T protein [Mus musculus]	Contains protein domain (PF00625) - Guanylate Kinase	UNCLASSIFIED	264685
2814	87643948 (5227, 5228)				22278996, 22278999, 29331825, 264508, 264906, 21906754, 264602, 264786, 264789, 52644229, 21906765, 33657109, 27488284, 18108370, 263972, 264555, 60432113
2815	87381996 (5229, 5230)			UNCLASSIFIED	264786, 18108394, 264682, 264693, 264508, 264509, 264907, 264626, 264908, 264909, 18108377, 264511, 264512, 264910, 264635, 264595, 265010, 264404, 264563, 264764, 264685, 264788

2616	87428895 (5231, 5232)	Novel Protein sim. GBank gij387676[emb]CAA828941 - (268780) predicted using GeneFinder. Similarity to Mouse FKBP-type binding protein (SW:FKBP_MOUSE) [Caenorhabditis elegans]	Contains protein domain (PF00234) - FKBP-type peptidyl-prolyl cis-trans isomerases	isomerase	22278995, 22278997, 22278998, 60433049, 60432289, 264828, 60433358, 264594, 60433438, 33109554, 87188474, 265011, 265017, 265019, 264288, 264768, 21908765, 21906767, 265020, 265021, 16106376, 16108377, 16108387, 87168518, 264482, 264567
2617	86978888 (5233, 5234)	Novel Protein sim. GBank gij728631[sp]P39188[ALU]_HUMAN - III ALU SUBFAMILY J WARNING ENTRY III		kinase	265010, 265018, 264369, 264693, 55611576, 22278002
2618	91231862 (5235, 5236)	Novel Protein sim. GBank gij318282 (AF049103) - Huntingtin interacting protein [Homo sapiens]	Contains protein domain (PF00397) - WW domain	UNCLASSIFIED	264489, 22278996, 264480, 264239, 29331822, 264102, 264509, 264908, 264907, 68712502, 29331830, 265008, 264910, 265009, 60433358, 60433438, 264758, 21906754, 265011, 87168559, 265017, 265018, 264369, 264288, 264768, 264768, 264689, 21906765, 21906768, 21906787, 33695917, 265020, 265022, 33657023, 264892, 33657109, 264626, 16106374, 33695855, 16108381, 83373044, 16108385, 16108388, 56528488, 264563
2619	87694000 (5237, 5238)	Novel Protein sim. GBank gij2431772 (U66411) - putative type III alcohol dehydrogenase [Drosophila melanogaster]	Contains protein domain (PF00465) - Iron-containing alcohol dehydrogenases	dehydrogenase	264259, 60432289, 60433438, 21906754, 264369, 60432113, 264566
2620	95314841 (5239, 5240)	Novel Protein sim. GBank gij322567[gi]AAD160971 - (AF090438) dactshund variant 1 [Mus musculus]	Contains protein domain (PF00828) - PHD-finger	UNCLASSIFIED	52644507, 52645156, 52646842, 65274572, 22278995, 56994075, 35698286, 22278999, 60432049, 264259, 52645080, 29331822, 29331824, 29331825, 29331828, 29331828, 35698052, 264907, 66712502, 265008, 60433358, 33657402, 52646317, 21906754, 87168474, 265010, 265017, 265018, 264448, 264369, 264288, 264687, 264768, 52644229, 264688, 264689, 21908785, 21906788, 35695917, 52644150, 264692, 33657108, 35695763, 35696423, 264556, 52644332, 16108382, 83373044, 16108385, 16108387, 65274727, 87168518, 60432113, 22279002
2621	80253495 (5241, 5242)	Novel Protein sim. GBank gij455734[trf]NP_001174, 11pA176 - ATPase, H+ transporting, lysosomal subunit 1; vacuolar proton pump; H+ ATPase subunit			264594, 264638
2622	81780380 (5243, 5244)	Novel Protein sim. GBank gij388035[emb]CAB052891 - (Z62285) predicted using GeneFinder [Caenorhabditis elegans]		UNCLASSIFIED	29331824, 35696052, 265007, 265010, 264288, 29148629

2624	91639308 (5247, 5248)	Novel Protein sim. GBank gij3880355[emb CA605299] - (262285) predicted using GeneFinder (Caenorhabditis elegans)		UNCLASSIFIED	56161688, 22276998, 22276997, 22276998, 22276999, 264259, 29331622, 29331624, 56162161, 29331625, 60432289, 29331626, 35696052, 29146489, 66712502, 52644045, 285007, 265006, 60433358, 33109954, 21908754, 265010, 265011, 265019, 264448, 264288, 21908765, 21908766, 21908767, 29146628, 35695917, 265021, 265022, 27466265, 16106370, 60431528, 65811576, 35695855, 56162323, 16106385, 67168518, 22279002, 16108391
2625	86452068 (5248, 5250)	Novel Protein sim. GBank gij2867429[db AA24657] - (AB007687) KIAA0427 (Homo sapiens)		UNCLASSIFIED	264091, 264511, 263361
2626	16533797 (5251, 5252)	Novel Protein sim. GBank gij467416 (L20302) - actin filament protein (Gallus gallus)		struct	265008
2627	87636823 (5253, 5254)	Novel Protein sim. GBank gij68462[db A27307] - proline-rich phosphoprotein (gene PRH1, DB allele) - human		UNCLASSIFIED	22276998, 265007, 265009, 264446, 21906767, 265021, 264558, 67168516
2628	94846254 (5255, 5256)	Novel Protein sim. GBank gij3123552[emb CA118609] - (AL022578) dJ393P12.2 (hypothetical) Proline-rich protein KIAA0268 (Homo sapiens)		UNCLASSIFIED	22276997, 22276999, 264259, 60432049, 29331622, 22276999, 264259, 29331625, 29331627, 35696052, 29331626, 264907, 264909, 265008, 264591, 60433358, 60433438, 265010, 265017, 265016, 264368, 264268, 16108357, 21908765, 21908766, 265022, 65274791, 264638, 16108387, 67168518, 22279002
2629	67376490 (5257, 5258)	Novel Protein sim. GBank gij4928585[db AA034056.1]AF15162 - (AF151621) C6H53 protein (Homo sapiens)		synthase	29331625, 29331628, 264102, 265006, 264766, 35695917, 264691, 33657023, 263972, 16108374, 22279000
2630	78186364 (5259, 5260)	Novel Protein sim. GBank gij321605[db Q1181] - Gag protein - Virus virus (strain EV1)	Contains protein domain (PF00086) - Zinc finger, CCHC class	UNCLASSIFIED	264636, 16108385
2631	94845909 (5261, 5262)	Novel Protein sim. GBank gij321605[db Q1181] - Gag protein - Virus virus (strain EV1)		dna_ma_bind	52644507, 52645156, 52646365, 52646842, 22276994, 22276995, 35696286, 56994075, 22276997, 22276999, 264259, 52645080, 29331622, 29331624, 29331625, 29331626, 29331627, 29331628, 35696052, 35696970, 264905, 264508, 264907, 264908, 264511, 264512, 265007, 265008, 264910, 52646317, 33657064, 52644286, 265010, 67168559, 265017, 265018, 265019, 264780, 264782, 264446, 264266, 264368, 264786, 264788, 52644229, 21906784, 21906765, 21906766, 21906787, 21906789, 35695917, 265020, 52644150, 33657023, 52645129, 33657109, 33657162, 27466261, 27466262, 27466265, 33657348, 35695763, 35696423, 65274791, 35695655, 264634, 264637, 52644332, 56162323, 60432113, 264588, 264488
2632	36730414 (5263, 5264)				264885

2633	95011617 (5265, 5266)	Novel Protein sim. Gbank gij13954dijibjBAA108891 - (DBA009) seizure-related gene product 6 type 2 precursor (Mus musculus)	Contains protein domain (PF00064) - Sushi domain (SCR repeat)		22278995, 22278997, 22278998, 264259, 29331822, 29331824, 29331825, 29331827, 264508, 264908, 265008, 265007, 265009, 265009, 55812038, 33657084, 55811386, 265010, 265011, 87168559, 265016, 265019, 264683, 264288, 264686, 28146829, 33657023, 264893, 33857182, 35895763, 55811576, 264639, 56162323, 63373044, 18108385, 56526486, 87168518, 22279000, 22279002, 264585
2634	87330921 (5287, 5288)	Novel Protein sim. Gbank gij5441611embjCAB4884.11 - (AJ38555) hypothetical protein [Canis familiaris]		UNCLASSIFIED	29331826, 263972, 264089
2635	86623144 (5289, 5270)	Novel Protein sim. Gbank g14680863jgJACD2721.11AF13294 - (AF132946) CGI-12 protein [Homo sapiens]			22278997, 264259, 29331824, 66714117, 29331827, 29331828, 264907, 33657084, 265017, 265018, 264448, 264288, 21906766, 21908767, 21906768, 29146829, 18108376, 55811576, 35895855, 87168518, 22279000
2636	87280534 (5271, 5272)	Novel Protein sim. Gbank gij3879146jembjCAB076461 - (Z83386) Similarly to Yeast hypothetical 52.9 KD protein (SW:P43619); cDNA EST EMBL:M69432 comes from this gene; cDNA EST EMBL:071008 comes from this gene; cDNA EST EMBL:073578 comes from this gene; cDNA EST EMBL:068025 come ...	Contains protein domain (PF01546) - Peptidase family M20/M23/M40		264569, 29331822, 29331828, 265008, 60170831, 264661, 264765, 264865, 29148627, 21906769, 29148764, 285022, 60170815, 284835, 18108385, 56526486, 22279002, 264567
2637	95011299 (5273, 5274)	Novel Protein sim. Gbank gij4758208jennjP_004081.11pDUSP - dual specificity phosphatase 3 (varicella virus phosphatase VH1-related)	Contains protein domain (PF00782) - Dual specificity phosphatase, catalytic domain	phosphatase	264468, 264489, 52844507, 264887, 52646365, 52846842, 22278994, 22278996, 22278999, 20281171, 284259, 29331822, 52645080, 66714117, 29331825, 29331826, 29331827, 35896052, 29331828, 29146498, 264508, 264905, 264908, 264907, 264908, 264909, 264511, 264512, 264910, 264591, 264592, 60432229, 264593, 284594, 33657402, 60433356, 264757, 60433438, 284598, 264758, 52846317, 21906754, 52844296, 285010, 264600, 264602, 264603, 264605, 264761, 264762, 264681, 264448, 264764, 264765, 264288, 264768, 264886, 264768, 264687, 264769, 21908785, 21906768, 21908787, 21908788, 21906769, 35895917, 285021, 60170815, 264891, 33657023, 264692, 264693, 65274620, 27486264, 18108370, 264628, 284829, 16108374, 35898423, 35895855, 264632, 264634, 264635, 264637, 264638, 52644332, 264639, 264558, 83373044, 60432113, 264594, 264595, 264596, 264488, 264587

2638	84326733 (5275, 5278)	Novel Protein sim. GBank gll4929889jgblAAC34105.1(AE15186 - (AE15188) CGI-110 protein [Homo sapiens])	Contains protein domain (PF00076) - RNA recognition motif, (a.k.a. RRM, RBD, or RNP domain)	UNCLASSIFIED	60424179, 52644507, 52646842, 18108399, 56182575, 22276995, 22276996, 35698286, 22276997, 22276999, 264259, 60432049, 29331822, 80424289, 29331828, 35696052, 29146498, 264905, 52644045, 56182435, 60433356, 35657402, 55612038, 55611386, 285019, 264288, 264789, 52644229, 58181562, 29148627, 29148629, 55811957, 29148794, 35695917, 265021, 52644150, 33657023, 85274820, 33657109, 35695763, 18108374, 55610784, 35696423, 55811578, 35695655, 60431850, 58182323, 60432113, 264404
2639	95381346 (5277, 5278)	Novel Protein sim. GBank gll2190007jgblBAAC03551 - (AB004109) phosphatidylserine synthase II [Cricetulus griseus]		synthase	264498, 29331825, 35696052, 264508, 264509, 264909, 264512, 33657402, 60433438, 284758, 85658542, 264600, 285020, 265021, 33657109, 264828, 35696423, 264555, 264639, 264563, 264564, 284585, 264598, 264488
2640	87781330 (5279, 5280)	Novel Protein sim. GBank gll3158516 (AF087617) - contains similarity to chromo (chromatin organization modifier) domains (Plant: chrom.hmm, score: 17.76 and 27.94) and to helicases conserved C-terminal domain (Plant: helicase_C.hmm, score: 87.00) [Caenorhabditis elegans]	Contains protein domain (PF00271) - Helicases conserved C-terminal domain	helicase	29331822, 29331826, 264908, 33109954, 285017, 265019, 21908788, 35695783, 264836, 264837, 18108387
2641	11869634 (5281, 5282)	Novel Protein sim. GBank gll2564955 (AF030001) - unknown [Mus musculus]		UNCLASSIFIED	264828
2642	87412875 (5283, 5284)	Novel Protein sim. GBank gll2564955 (AF030001) - unknown [Mus musculus]		UNCLASSIFIED	264259, 29331822, 33657402, 265019, 264369, 264691, 264834, 56526486, 22279002
2643	87643961 (5285, 5286)	Novel Protein sim. GBank gll4490304jemb[CAB36795.1 - (AL035678) putative protein [Arabidopsis thaliana]	Contains protein domain (PF00270) - DEAD/DEAH box helicase	helicase	22276997, 264259, 29148499, 56182435, 264910, 265010, 18108351, 264892, 264893, 264369, 264694, 264895, 264896, 29148827, 264690, 33657109, 18108370, 263973, 18108374, 264834, 264557, 264556, 18108385, 264482
2644	88177671 (5287, 5288)	Novel Protein sim. GBank gll3789797jgblAAC67502.1 - (AF059568) actin binding protein MAYVEN [Homo sapiens]	Contains protein domain (PF00651) - BTB/POZ domain	nuc_rept	284107, 264887
2645	17271228 (5289, 5290)	Novel Protein sim. GBank gll1706722jgblP49749IEVX2 - MOUSE - HOMEBOX EVEN SKIPPED HOMOLOG PROTEIN 2 (EVX2)		UNCLASSIFIED	265007
2646	84148542 (5291, 5292)	Novel Protein sim. GBank gll1706722jgblP49749IEVX2 - MOUSE - HOMEBOX EVEN SKIPPED HOMOLOG PROTEIN 2 (EVX2)		UNCLASSIFIED	264909, 264687, 264632, 83373044

2847	81212876 (5283, 5284)			UNCLASSIFIED	5618257, 22278998, 35696288, 22278998, 284259, 28331822, 56182181, 28331825, 60424289, 60432289, 35696052, 68712502, 284908, 285007, 55812038, 33109954, 21906754, 33657084, 285019, 284448, 284288, 56181562, 21906765, 21906766, 21906768, 21906769, 35695917, 285020, 285021, 52844150, 284693, 33657109, 33657349, 60431528, 18108374, 55610764, 35696423, 58182323, 60432113, 22278902, 284564
2848	67600567 (5295, 5296)				28146498, 56182435, 33109954, 265011, 284882, 55811957, 35695917, 284890, 263976, 18108377, 35698423, 80432113, 56182575, 35696288, 22278998, 28331824, 28331828, 60432229, 33657402, 33109954, 60170831, 60432229, 33657402, 33109954, 21906754, 285017, 284886, 284688, 21906785, 21906766, 60170815, 284693, 263987, 18108370, 283978, 60170394, 60432113, 22279002, 284563
2849	84126783 (5297, 5298)	Novel Protein sim. GBank gij3041652 (AC004539) - unknown function; similar to Y08105 (P/D;g1686171) [Homo sapiens]		UNCLASSIFIED	29331824, 29331826, 29331827, 265007, 55812038, 21906754, 18108366, 18108384, 22278902, 264567
2850	67287533 (5289, 5300)	Novel Protein sim. GBank gij5380271 (dbj BAA81908.1) - (AB029335) HPE1-3 (Haloecynthia cretzi)			264892
2851	88086745 (5301, 5302)	Novel Protein sim. GBank gij4240225 (dbj BAA74691.1) - (AB020675) KIAA0868 protein (Homo sapiens)	Contains protein domain (PF00054) - synthase lamelin G domain		264892
2852	10343125 (5303, 5304)	Novel Protein sim. GBank gij4483958 (emj CA811123.2) - (Z98551) predicted using hexExon; MAL3P6.28 (PF008450), Hypothetical protein, len. 187 aa. Similarly to model organism hypothetical proteins (C.elegans, D.melanogaster, S.cerevisiae & S.pombe), C.elegans protein ZK287.5 (Tr....)		UNCLASSIFIED	265018, 18108370, 18108387, 264566
2853	87788735 (5305, 5306)				
2854	85103240 (5307, 5308)				60424178, 65274572, 56182575, 284259, 56182181, 284808, 56182435, 55811957, 35695917, 265021, 283978, 55810764, 65274791, 56182323, 63373044, 65274127
2855	81228018 (5309, 5310)	Novel Protein sim. GBank gij3873272 (emj CA802881.1) - (Z81051) predicted using GeneFinder; similar to Zinc finger, C3HC4 type (RING finger); cDNA EST yk443h5.3 comes from this gene; cDNA EST yk443h5.5 comes from this gene [Caenorhabditis elegans]	Contains protein domain (PF00097) - Zinc finger, C3HC4 type (RING finger)	transcriptase	56182575, 56181886, 284092, 284259, 56182181, 60432288, 264807, 33657402, 55812038, 21906754, 67168558, 265017, 284446, 284389, 284286, 21906765, 21906766, 21906767, 21906768, 33657109, 16108370, 264628, 55811576, 264558, 284639, 63373044, 58528486, 284404, 60432113
2856	84582601 (5311, 5312)	Novel Protein sim. GBank gij3043718 (dbj BAA25523) - (AB011189) KIAA0587 protein (Homo sapiens)			284883

2657	52561728 (5313, 5314)	Novel Protein sim. GBank gll5685091dbj BAAG3038.11 - (AB028009) KIAA1088 protein [Homo sapiens]		dna_mn_bind	264693
2656	88062454 (5315, 5316)	Novel Protein sim. GBank gll5685088 (AC005757) - R32611_1 [Homo sapiens]	Contains protein domain (PF00560) - nucleosidehnb Leucine Rich Repeat		35696286, 264259, 29331622, 28331624, 29331826, 29331828, 265018, 264683, 21906768, 35695917, 264693, 35695855, 264637, 87168518, 284486, 264367
2659	87600755 (5317, 5318)	Novel Protein sim. GBank gll54203871emb CA846876.11 - (AJ234559) proteophosphoplycan [Leishmania major]	Contains protein domain (PF01426) - BAH domain Contains protein domain (PF00036) - EF hand	UNCLASSIFIED	264908, 264910, 265018, 264369, 264769, 21906769, 264693, 265972, 16108386
2660	91718472 (5319, 5320)	Novel Protein sim. GBank gll7288375spl P9194/ALU1_HUMAN - III ALU SUBFAMILY SQ WARNING ENTRY III		kinase	264488, 65274572, 35696286, 22278998, 22278999, 264259, 29331622, 28331824, 60432289, 29331826, 35696052, 264808, 56182435, 265008, 265009, 60433358, 264594, 265010, 265018, 55811150, 18108351, 264682, 264684, 264369, 264288, 264687, 21906765, 29148784, 35695917, 60170615, 52644150, 33657023, 33657109, 35696423, 35695855, 264556, 60170394, 16108385, 22279000, 22279002
2661	95342817 (5321, 5322)	Novel Protein sim. GBank gll4758048p NP_004739.1pCPR8 - cell cycle progression 8 protein		glycoprotein	60432049, 264259, 29331824, 29331825, 29331826, 29331827, 29331828, 264906, 264809, 264593, 33108954, 265010, 265017, 265018, 265019, 264760, 264448, 264369, 264288, 21906765, 21906768, 265022, 264691, 33657023, 27488282, 80431528, 16108374, 35695855, 18108388, 264482
2662	80228739 (5323, 5324)	Novel Protein sim. GBank gll3874714emb CA912631 - (Z68494) similar to choline dehydrogenase; cDNA EST yk346d5.5 comes from this gene; cDNA EST yk346d5.3 comes from this gene [Caenorhabditis elegans]		dehydrogenase	264555, 264556, 264558, 264486
2663	87780823 (5325, 5326)	Novel Protein sim. GBank gll3874714emb CA912631 - (Z68494) similar to choline dehydrogenase; cDNA EST yk346d5.5 comes from this gene; cDNA EST yk346d5.3 comes from this gene [Caenorhabditis elegans]			264908, 264909, 264757, 264758, 264787, 264691, 33657023, 284638
2664	85518328 (5327, 5328)	Novel Protein sim. GBank gll3388670 (U56977) - Notch homolog Scalloped wings [Lucilia cupulina]	Contains protein domain (PF00008) - EGF-like domain	oncogene	35696288, 264509, 264595, 264268, 264895, 264888
2665	87770682 (5329, 5330)	Novel Protein sim. GBank gll4884081emb CA843311.11 - (AL050190) hypothetical protein [Homo sapiens]		UNCLASSIFIED	35696286, 22278998, 29331822, 35696052, 264906, 264907, 264809, 264510, 264511, 264512, 264593, 60433438, 265019, 264881, 21906765, 21906766, 21906767, 21906768, 265020, 265022, 35696423, 35695855, 22279002, 264482, 264488
2666	87828472 (5331, 5332)	Novel Protein sim. GBank gll5106958p AACD39808.11AF113615 - (AF113615) FH1/FH2 domain-containing protein FHO5 [Homo sapiens]		UNCLASSIFIED	29331825, 265007, 264910, 60432222, 265019, 264288, 21908767, 264558, 22279002
2667	87422720 (5333, 5334)	Novel Protein sim. GBank gll2500570sp Q17533 RNPH_CAEEL - RIBONUCLEASE PH-LIKE PROTEIN B0564.1	Contains protein domain (PF01138) - nuclease 3' exonuclease family	nuclease	264807, 29331830, 264881, 264883, 264288, 35695855, 264832, 264556, 264557, 264558, 264559, 264583, 264565, 264567

2866	91218718 (5335, 5336)	Novel Protein sim. GBank gil5454186 nfp_008327.1 p27G1 - ZYG homolog		UNCLASSIFIED	56181688, 35696286, 22278998, 22278999, 56182181, 29331824, 60424289, 29331825, 35698052, 29331828, 68712502, 56182435, 60433358, 264758, 21906754, 55811386, 265011, 87168559, 265017, 265019, 55811150, 264448, 264389, 264288, 21906785, 21908766, 21808768, 55811957, 35695917, 265020, 265021, 33657023, 284892, 33657109, 35695763, 60431528, 18108374, 35696423, 55811576, 264834, 60431850, 83373044, 18108385, 87188518, 22279000, 264563, 264564
2689	95415721 (5337, 5338)	Novel Protein sim. GBank gil2147012 p1 JC4899 - proline rich protein - rel			264489, 264889, 21906767, 65274572, 56182575, 21906766, 29148927, 21906769, 29148929, 35698286, 35695917, 22278998, 22278999, 265021, 265022, 80170615, 52644150, 60432049, 264258, 264891, 33657023, 264692, 29331822, 29331824, 29331825, 60432289, 33657109, 29331828, 29331827, 35698052, 29331828, 29148498, 29148489, 264905, 264908, 52644045, 29148489, 264905, 264908, 52644045, 35695895, 265006, 264910, 264835, 60432229, 264592, 264636, 58182323, 60433356, 60170394, 60433438, 264559, 264595, 55812038, 33109944, 87188559, 60432113, 265019, 264448, 264369, 264684, 264288
2670	87613234 (5339, 5340)	Novel Protein sim. GBank gil1723523 p9 C10382 YDBB_SCHPO - HYPOTHETICAL 94.9 KD PROTEIN C22E12.11C IN CHROMOSOME I	Contains protein domain (PF00828) - PHD-finger		16108370, 263974
2671	91214938 (5341, 5342)	Novel Protein sim. GBank gil4768277 p9 A029444.1 AF06425 - (AF06425) very long-chain acyl-CoA synthetase homolog 2. VLCS-H2 (Homo sapiens)		transport	52646842, 56994075, 264259, 29331822, 29331824, 29331825, 29331827, 33658970, 264509, 265006, 33109954, 21906754, 264882, 264288, 265021, 33657023, 33657109, 33657182, 27488261, 27488282, 27486285, 18108376, 18108385
2672	87399123 (5343, 5344)	Novel Protein sim. GBank gil4986348 p9 A034877.1 AC00834 - (AC00834) Contains two PF01344 Kelch motif domains. [Arabidopsis thaliana]	Contains protein domain (PF01344) - Kelch motif	UNCLASSIFIED	264787

2873	07430748 (5345, 5346)	Novel Protein sim. GBank gll5457337[emb]CAB4.1505.21 - (A1238876) poly(ADP-ribose) polymerase 2 [Homo sapiens]	Contains protein domain (PF00644) - Poly(ADP-ribose) polymerase catalytic region.	polymerase	22278995, 22278996, 22278997, 22278999, 29331822, 29331824, 29331828, 35696052, 80433438, 87168474, 87168558, 285017, 265018, 285019, 284448, 21906768, 21906769, 285020, 265021, 33657109, 27486262, 35695763, 60431850, 60170394, 87168518, 284563
2874	84047721 (5347, 5348)	Novel Protein sim. GBank gll4758824[ref]NP_004280.1[pNRF3 - nuclear factor (erythroid-derived 2)-like 3	Contains protein domain (PF00170) - bZIP transcription factor	transcription factor	284408, 22278996, 35696266, 264091, 264259, 29331824, 29331828, 35696052, 284511, 55812038, 85658542, 264766, 21906765, 35695917, 264628, 35694423, 18108383, 87168518
2875	79563835 (5349, 5350)			UNCLASSIFIED	264906, 265008
2876	79828383 (5351, 5352)			UNCLASSIFIED	18108394, 85274572, 56182575, 22278994, 22278995, 56894075, 22278998, 35696288, 22278997, 22278999, 284490, 264259, 52845080, 29331824, 29331825, 60432288, 29331827, 29331828, 35696052, 28148499, 29331830, 284908, 52844045, 265008, 265007, 265008, 265009, 60432228, 60433396, 80433438, 55812038, 265010, 265011, 87168558, 265017, 265019, 18108351, 264682, 284448, 284883, 284288, 21906785, 21906766, 21908787, 21906768, 21908769, 55811957, 265020, 265021, 60170615, 52844150, 284691, 33657023, 263967, 33657109, 27486284, 27486285, 33657349, 35695763, 18108370, 18108374, 18108377, 55811576, 35696423, 35695855, 83373044, 18108387, 22279000, 22279002, 284594
2877	84328800 (5353, 5354)	Novel Protein sim. GBank gll1078042[pH]SS2154 - acetyl-CoA synthetase - fruit fly (Drosophila melanogaster)	Contains protein domain (PF00501) - AMP-binding enzyme	synthase	18108394, 85274572, 56182575, 22278994, 22278995, 56894075, 22278998, 35696288, 22278997, 22278999, 284490, 264259, 52845080, 29331824, 29331825, 60432288, 29331827, 29331828, 35696052, 28148499, 29331830, 284908, 52844045, 265008, 265007, 265008, 265009, 60432228, 60433396, 80433438, 55812038, 265010, 265011, 87168558, 265017, 265019, 18108351, 264682, 284448, 284883, 284288, 21906785, 21906766, 21908787, 21906768, 21908769, 55811957, 265020, 265021, 60170615, 52844150, 284691, 33657023, 263967, 33657109, 27486284, 27486285, 33657349, 35695763, 18108370, 18108374, 18108377, 55811576, 35696423, 35695855, 83373044, 18108387, 22279000, 22279002, 284594

2678	95001684 (5355, 5356)	Novel Protein sim. GBank gll66760pH4A0465 - alanine transaminase (EC 2.6.1.2), cytosolic - human		UNCLASSIFIED	264468, 263994, 264489, 16108394, 52846842, 35696286, 22276998, 264259, 29331825, 35696052, 264508, 264905, 264509, 264906, 264907, 264908, 264909, 264510, 264511, 265006, 264512, 265007, 265008, 264910, 265009, 264593, 264594, 264598, 264758, 33108954, 52646317, 21906754, 265010, 265011, 87168559, 264600, 264601, 264602, 264603, 265017, 265019, 264605, 264760, 264762, 16108351, 264763, 264682, 264683, 264764, 264286, 264768, 264687, 264768, 264769, 264689, 21908765, 21908767, 35695917, 285020, 265021, 52644150, 264691, 33657023, 264693, 16108384, 16108385, 33657109, 33857349, 264828, 16108374, 35688423, 35695855, 264630, 264631, 264632, 264634, 264635, 264636, 264637, 264638, 264639, 16108385, 264563, 264564, 264565, 264586, 264488, 264587
2679	95361544 (5357, 5358)	Novel Protein sim. GBank gll709233spIP07514INCsr - NADH- CYTOCHROME B5 REDUCTASE	Contains protein domain (PF00970) - FAD/NAD-binding Cytochrome reductase	reductase	264488, 22278998, 35696286, 264259, 29331828, 29331827, 29331828, 264908, 56182435, 264113, 264511, 265008, 60433438, 264758, 65658542, 87188474, 265011, 265017, 265019, 264288, 21908766, 21908767, 21908768, 21908769, 55811957, 265021, 264690, 33657023, 55810764, 35696423, 55811576, 264631, 16108381, 60170394, 83373044, 87168518, 264586 264258, 264102, 264805, 264908, 265007, 265008, 33109954, 265011, 16108351, 264768, 33657023, 20281149, 263972, 264630, 264635, 264838
2680	87800356 (5359, 5360)	Novel Protein sim. GBank gll4589604idejBAA76824.11 - (AB023197) KIAA0980 protein [Homo sapiens]	Contains protein domain (PF00036) - EF hand	struct	264489, 22278995, 264508, 264905, 264908, 264907, 264908, 264909, 264510, 265006, 264512, 264910, 264594, 80433436, 264758, 264603, 264604, 265019, 264605, 264760, 264764, 264687, 264768, 264769, 21908769, 35695917, 265020, 33657023, 264631, 264835, 264837, 264838, 264639, 264486
2681	80933844 (5361, 5362)	Novel Protein sim. GBank gll728837spIP39184ALUT - HUMAN - III ALU SUBFAMILY SO WARNING ENTRY IIII		kinase	35696286, 264808, 55811366, 265017, 55811550, 55811957, 35695917, 80431528, 55810764, 55811576, 35696423, 65274791, 56526486
2682	94138934 (5363, 5364)	Novel Protein sim. GBank gll423468pH4J01974 - HTF9-C protein - mouse	Contains protein domain (PF00078) - RNA recognition motif (a.k.a. RRM, RBD, or RNP domain)	UNCLASSIFIED	
2683	87774405 (5365, 5366)	Novel Protein sim. GBank gll51143519pJAA040286.11 - (AF156271) RING finger protein tierf [Homo sapiens]	Contains protein domain (PF00622) - SPRY domain	interleukinrecept	

2684	85767151 (5387, 5388)	Novel Protein sim. GBank g 14886469emb CAB3385.11 - (AL050284) hypothetical protein [Homo sapiens]			264593
2685	88054298 (5369, 5370)	Novel Protein sim. GBank g 3342729 (AC005331) - R31341.2 [Homo sapiens]		UNCLASSIFIED	
2686	87628680 (5371, 5372)	Novel Protein sim. GBank g 14850844 cb BAA77027.11 - (AB026190) Kelch motif containing protein [Homo sapiens]	Contains protein domain (PF01344) - Kelch motif	struct	264258, 29331822, 60432289, 29331827, 29331830, 284908, 284512, 284596, 284789, 284534, 264555, 264556, 264557, 264558, 80170394, 264559, 264488
2687	87986183 (5373, 5374)	Novel Protein sim. GBank g 15281314 gb AA041475.1 AF133123 transcription factor IIC102 [Homo sapiens]	Contains protein domain (PF00515) - TPR Domain	transcriptfactor	18108394, 18108396, 22278996, 35696286, 22278997, 29331826, 29331828, 86712502, 21906754, 285011, 264780, 264781, 264783, 264688, 21906765, 35696423, 264559, 18108385, 264563
2688	78959584 (5375, 5376)				264908, 264760
2689	94122440 (5377, 5378)	Novel Protein sim. GBank g 3880023 emb CAA873391 - (Z73098) Similarly to yeast hypothetical protein [Swiss Prot accession number Q08895]. cDNA EST EMBL:D72982 comes from this gene; cDNA EST EMBL:D75947 comes from this gene [Caenorhabditis elegans]	Contains protein domain (PF00560) - Leucine Rich Repeat		22278897, 22278998, 22278999, 29331824, 35698052, 284908, 284909, 56182435, 264512, 264910, 265009, 60433438, 21906754, 18108351, 264682, 264683, 264787, 21906765, 21906768, 21906769, 33657023, 33657182, 27486262, 27486264, 33657349, 18108370, 18108372, 18108374, 35696423, 35695855, 18108385, 22279002
2690	88003055 (5379, 5380)	Novel Protein sim. GBank g 2477513 (AC002386) - F25965.3 [Homo sapiens]	Contains protein domain (PF00620) - RhoGAP domain	struct	
2691	81218241 (5381, 5382)	Novel Protein sim. GBank g 416727 emb CAA671301 - (X98506) acetyl-CoA synthetase [Solanium tuberosum]		synthase	65274572, 38182575, 264259, 29331822, 29331824, 29331826, 29331828, 264112, 265009, 55812038, 264598, 33109954, 265017, 264448, 264288, 264358, 284684, 21906769, 80170815, 60431528, 55810764, 264634, 264636, 264556, 264637, 22279002, 264564, 264568
2692	94111914 (5383, 5384)	Novel Protein sim. GBank g 3513303 (AC005594) - R26984.1 [Homo sapiens]	Contains protein domain (PF00326) - Proyl oligopeptidase family	peptidase	
2693	20438807 (5385, 5386)	Novel Protein sim. GBank g 3122400 sp O35862 MUJ_MOUSE - MYELOID UPREGULATED PROTEIN		UNCLASSIFIED	264592
2694	94111916 (5387, 5388)				264559
2695	95345513 (5389, 5390)	Novel Protein sim. GBank g 1487274 gb AA034785.11 - (AF132177) unknown [Drosophila melanogaster]		collagen	35686286, 56984075, 22278998, 284258, 35686052, 29331830, 265011, 264288, 56181562, 264690, 264692, 33657023, 27486262, 263876, 18108376, 35698423, 35698855, 80170394, 83373044, 56526486, 22279000, 22279002, 264588
2696	87874040 (5391, 5392)	Novel Protein sim. GBank g 1728831 sp P38186 ALU1_HUMAN - III ALU SUBFAMILY J WARNING ENTRY III		synthase	264594, 21906768, 18108370, 18108372

2697	97636472 (5393, 5394)	Novel Protein sim. GBank glij5689473dijl5AA63020.1 - (AB028991) KIAA1088 protein [Homo sapiens]		UNCLASSIFIED	35698288, 264259, 29331824, 29331825, 29331828, 29331827, 29331828, 35698052, 265006, 265007, 284512, 285009, 60170831, 60433358, 264585, 264758, 87186474, 265010, 285011, 87188558, 264601, 265017, 265018, 285019, 264761, 18108351, 264448, 264288, 264768, 264688, 264689, 21908765, 21908768, 265020, 265021, 80170815, 33857109, 18108376, 35698423, 35698555, 264555, 264558, 60170394, 284559, 18108387, 56526486, 67168518, 22279002, 264583, 264482
2698	84325891 (5395, 5396)	Novel Protein sim. GBank glij841318 (UJ22818) - mutant sterol regulatory element binding protein-2 [Oriculus griseus]	Contains protein domain (PF00412) - LIM domain containing proteins	transcriptfactor	22278995, 35686288, 22278998, 22278997, 22278998, 22278999, 284259, 29331822, 29331824, 29331825, 60432289, 29331828, 29331827, 33656870, 264806, 29331830, 284909, 52644045, 264810, 60433356, 33657402, 33109954, 265017, 285018, 265019, 264288, 21908765, 21906766, 21906767, 21906768, 29148629, 35685817, 285021, 285022, 52844150, 33857023, 33657182, 27486281, 35698423, 65274791, 284638, 60432113, 22279000
2699	87780650 (5397, 5398)	Novel Protein sim. GBank		UNCLASSIFIED	264768, 18108357, 284690, 264691
2700	84139838 (5399, 5400)	Novel Protein sim. GBank glij5174395/ejlfP_006006.1ipb120 - Brain protein 120		UNCLASSIFIED	80424289, 56182435, 60432229, 60433338, 55811386, 265017, 55811150, 264448, 56181582, 55811857, 264693, 33657109, 60431528, 264629, 55810764, 55811576, 65274791, 60431850, 60432113
2701	84148584 (5401, 5402)	Novel Protein sim. GBank glij572801 (U70854) - F38A5.1 gene product [Caenorhabditis elegans]			18108394, 52645156, 35686286, 264259, 29331822, 29331824, 29331825, 80432289, 29331828, 29331827, 35696052, 29331828, 29148499, 285008, 80433358, 33657402, 60433338, 264595, 33657084, 18108351, 264288, 264769, 18108359, 21906768, 35695917, 33657023, 27486261, 18108374, 18108379, 35696423, 18108382, 83373044, 18108384, 18108388, 60432113, 22279000
2702	57283368 (5403, 5404)	Novel Protein sim. GBank glij2605987 (AF030027) - 24 [Equine herpesvirus 4]		UNCLASSIFIED	29331828, 264512, 264555, 264556, 264559
2703	87649514 (5405, 5406)	Novel Protein sim. GBank glij5689399[dijl]BA62883.1 - (AB028954) KIAA1031 protein [Homo sapiens]	Contains protein domain (PF00642) - Zinc finger C-x8-C-x5-C-x3-H type (end similar)		80432288, 285007, 21906785, 21908768, 265021, 264553

2704	87649515 (5407, 5408)	Novel Protein sim. GBank gll335864 (gll3AAB63294) - (AF008554) Implantation-associated protein (Rattus norvegicus)			264488, 22278995, 22278998, 29331828, 29146489, 264905, 264806, 264907, 52844045, 264511, 33657402, 264600, 264802, 265017, 264605, 264781, 18108351, 264764, 264887, 264769, 265021, 264691, 264692, 18108362, 264693, 18108370, 18108374, 264834, 264835
2705	87771745 (5409, 5410)				264489, 264509, 264511, 264512, 264810, 264593, 87188474, 264604, 264298, 264887, 264759, 264638, 264568, 264466
2708	94328789 (5411, 5412)	Novel Protein sim. GBank gll3255852 (embhCA16821.1) - (AL021728) /prediction=(method: match)=(desc: Drosophila melanogaster)	UNCLASSIFIED		264488, 52646842, 65274572, 22278994, 58894075, 22278997, 264239, 29331824, 29331825, 29331826, 29331828, 33658970, 264807, 264908, 264909, 52844045, 56182435, 265006, 265007, 60433438, 55612036, 21906754, 52844298, 285010, 264801, 265017, 265018, 264681, 264448, 264882, 264268, 264888, 264887, 264888, 21906766, 21906769, 55611957, 35695917, 265020, 265021, 60170815, 264690, 264691, 33657023, 264692, 264693, 65274620, 27488284, 263972, 18108374, 18108377, 264635, 264638, 264558, 80170394
2707	86089639 (5413, 5414)	Novel Protein sim. GBank gll3417294 (AC004381) - Unknown gene product (Homo sapiens)			83373044, 65274727, 87169518, 22279000, 22278996, 22278998, 58182435, 21908754, 87168559, 265017, 264448, 52645129
2708	81011351 (5415, 5416)	Novel Protein sim. GBank gll545780 (bbsl147178 - DARPP-32-dopamine and cAMP-regulated phosphoprotein (human, brain, Peptide, 204 aa)	UNCLASSIFIED		65274572, 264239, 29331822, 29331823, 60432288, 29331826, 29331827, 29331828, 264909, 264510, 265007, 264910, 60433356, 60433438, 33108954, 265010, 285011, 264389, 264288, 264765, 264693, 264565
2709	84853888 (5417, 5418)	Novel Protein sim. GBank gll3169705 (AC004780) - F17127.1 (Homo sapiens)	UNCLASSIFIED		29331822, 16108370, 18108374, 83373044
2710	87627979 (5419, 5420)	Novel Protein sim. GBank gll446831 (embhCA837992) - (AL031432) d146SN24.2.1 (PUTATIVE novel protein) (isoform 1) (Homo sapiens)			29331824, 264759, 264693, 16108382, 16108386

2711	94111820 (5421, 5422)	Novel Protein sim. GBank gi13122400 pC35692 MLUG_MOUSE - MYELOID UPREGULATED PROTEIN		UNCLASSIFIED	264498, 264697, 52645156, 264769, 21906764, 21906765, 21906767, 21906768, 21909799, 55611957, 56994075, 22279997, 22278998, 265020, 265021, 264690, 264258, 264691, 264692, 33657023, 29331822, 29331824, 60424269, 29331828, 33657192, 29331827, 27489262, 33657349, 264509, 264905, 264907, 60431528, 264909, 264909, 55610764, 35898423, 95274791, 35695855, 265007, 264910, 60431850, 60432229, 264557, 264558, 55912036, 33109954, 19108395, 21906754, 33657084, 97186518, 97168474, 67169559, 60432113, 285017, 22279000, 265018, 285018, 22279002, 264760, 55911150, 264881, 19109351, 264555, 264784, 264566, 264299, 264766 264498, 35986288, 22278989, 264259, 29331824, 90432289, 35696052, 264509, 264908, 96712502, 52644045, 265006, 60432229, 33657402, 60433358, 285010, 265018, 16109351, 264881, 264299, 264695, 21906765, 21906766, 21906788, 21906789, 55611957, 35695917, 285020, 285021, 60170615, 284691, 264692, 33657023, 33657109, 33657192, 33957349, 19106370, 18109374, 35698423, 35995855, 264555, 52644332, 56182223, 97186519, 60432113 66714117, 264906, 264563
2712	94312071 (5423, 5424)	Novel Protein sim. GBank gi15081315 gbp AAD39343.1 AF076607 - (AF076607) prediabetic NOD sera-reactive autoantigen [Mus musculus]	Contains protein domain (PF00515) - Transferase TPR Domain		264498, 264697, 52645156, 264769, 21906764, 21906765, 21906767, 21906768, 21909799, 55611957, 56994075, 22279997, 22278998, 265020, 265021, 264690, 264258, 264691, 264692, 33657023, 29331822, 29331824, 60424269, 29331828, 33657192, 29331827, 27489262, 33657349, 264509, 264905, 264907, 60431528, 264909, 264909, 55610764, 35898423, 95274791, 35695855, 265007, 264910, 60431850, 60432229, 264557, 264558, 55912036, 33109954, 19108395, 21906754, 33657084, 97186518, 97168474, 67169559, 60432113, 285017, 22279000, 265018, 285018, 22279002, 264760, 55911150, 264881, 19109351, 264555, 264784, 264566, 264299, 264766 264498, 35986288, 22278989, 264259, 29331824, 90432289, 35696052, 264509, 264908, 96712502, 52644045, 265006, 60432229, 33657402, 60433358, 285010, 265018, 16109351, 264881, 264299, 264695, 21906765, 21906766, 21906788, 21906789, 55611957, 35695917, 285020, 285021, 60170615, 284691, 264692, 33657023, 33657109, 33657192, 33957349, 19106370, 18109374, 35698423, 35995855, 264555, 52644332, 56182223, 97186519, 60432113 66714117, 264906, 264563
2713	88003084 (5425, 5426)	Novel Protein sim. GBank gi2477513 (AC002398) - F25965_3 [Homo sapiens]		UNCLASSIFIED	264636 264508, 264905, 264907, 264908, 264909, 284910, 284758, 265011, 284780, 284784, 284298, 264768, 294768, 284638 284091, 264259, 29331822, 96714117, 284906, 264369, 294893, 264558, 264593 264593, 264558
2714	13529218 (5427, 5429)	Novel Protein sim. GBank gi14321868 p AD159971 - (AF067430) Smarcat1-related protein [Mus musculus]		UNCLASSIFIED	264636 264508, 264905, 264907, 264908, 264909, 284910, 284758, 265011, 284780, 284784, 284298, 264768, 294768, 284638 284091, 264259, 29331822, 96714117, 284906, 264369, 294893, 264558, 264593 264593, 264558
2715	94122454 (5429, 5430)	Novel Protein sim. GBank gi13327046 p BAA315911 - F25965_3 [Homo sapiens]		UNCLASSIFIED	264636 264508, 264905, 264907, 264908, 264909, 284910, 284758, 265011, 284780, 284784, 284298, 264768, 294768, 284638 284091, 264259, 29331822, 96714117, 284906, 264369, 294893, 264558, 264593 264593, 264558
2716	88003088 (5431, 5432)	Novel Protein sim. GBank gi2477513 (AC002398) - F25965_3 [Homo sapiens]		glycoprotein	264636 264508, 264905, 264907, 264908, 264909, 284910, 284758, 265011, 284780, 284784, 284298, 264768, 294768, 284638 284091, 264259, 29331822, 96714117, 284906, 264369, 294893, 264558, 264593 264593, 264558
2717	88007461 (5433, 5434)	Novel Protein sim. GBank gi13327046 p BAA315911 - (AB014516) KIAA0819 protein [Homo sapiens]			264636 264508, 264905, 264907, 264908, 264909, 284910, 284758, 265011, 284780, 284784, 284298, 264768, 294768, 284638 284091, 264259, 29331822, 96714117, 284906, 264369, 294893, 264558, 264593 264593, 264558
2718	78604062 (5435, 5436)	Novel Protein sim. GBank gi1746495 (U23515) - weakly similar to gastrin zinc finger protein [Caenorhabditis elegans]		UNCLASSIFIED	264636 264508, 264905, 264907, 264908, 264909, 284910, 284758, 265011, 284780, 284784, 284298, 264768, 294768, 284638 284091, 264259, 29331822, 96714117, 284906, 264369, 294893, 264558, 264593 264593, 264558
2719	88190423 (5437, 5439)				264636 264508, 264905, 264907, 264908, 264909, 284910, 284758, 265011, 284780, 284784, 284298, 264768, 294768, 284638 284091, 264259, 29331822, 96714117, 284906, 264369, 294893, 264558, 264593 264593, 264558

2720	95086242 (5439, 5440)	Novel Protein sim. GBank gl 335873 (U6889) - ATP-dependent RNA helicase [Mus musculus]	Contains protein domain (PF00270) - helicase DEAD/DEAH box helicase		18108374, 60424179, 264489, 56182435, 21906765, 21906766, 35689423, 22276997, 265020, 265022, 265006, 265008, 264092, 264636, 60432229, 264691, 264692, 33657023, 264693, 33657402, 63373044, 29331824, 18108366, 60424269, 28331826, 18108385, 52645129, 21906754, 35696052, 29331826, 87188474, 264100, 265010, 285011, 265019, 22279002, 264905, 264482, 264563, 264906, 18108351, 264681, 18108370, 29331630, 264908, 68712502, 5264045, 264909, 264628, 18108354
2721	95345523 (5441, 5442)	Novel Protein sim. GBank gl 4929683 gb AA034092.1 AF15185 - (AF15185) CGI-87 protein [Homo sapiens]	Contains protein domain (PF01172) - Uncharacterized protein family UFP0023		22278995, 35686286, 264259, 29331622, 29331824, 66714117, 29331826, 264906, 60433438, 265017, 18108351, 264448, 264288, 264769, 21906788, 265021, 33657109, 283969, 60431528, 264629, 55811576, 55274791, 35695855, 264631, 264637, 60170394, 56182223, 22279000
2722	81838807 (5443, 5444)	Novel Protein sim. GBank gl 3212897 gb AA023434.1 - (AC004997) match to ESTs AA687999 (NID:82626700), AA185485 (NID:81741481), Z45871 (NID:855105), and T84026 (NID:8712314), similar to various Tre-like proteins including: AF040854 (P D 82746883), D13644 (P D 82104571), AL02114...	Contains protein domain (PF00566) - oncogene TBC domain		35686286, 22278999, 21906754, 265017, 264782, 264288, 21906765, 21906767, 21906768, 35695917, 18108362, 27486262, 35695855, 264558, 264559
2723	87387732 (5445, 5446)			UNCLASSIFIED	284506, 264508, 264806, 264808, 264810, 55612036, 264786, 264687, 264628, 264636, 264486
2724	87639363 (5447, 5448)	Novel Protein sim. GBank gl 4680881 gb AA027730.1 AF13295 - (AF13295) CGI-21 protein [Homo sapiens]		ubiquitin	18108396, 22278998, 20281089, 28331824, 29331826, 60432229, 28331828, 60170831, 60432229, 60433436, 18108351, 264682, 21906766, 21906767, 21906769, 35695917, 33657023, 33657109, 18108372, 18108374, 35695855, 22279000, 22279002
2725	94853991 (5449, 5450)	Novel Protein sim. GBank gl 3169705 (AC004780) - F17127_1 [Homo sapiens]		UNCLASSIFIED	264486, 32644507, 264259, 28331627, 21906754, 285011, 18108351, 264448, 264286, 264665, 264689, 35695917, 265020, 33657182, 27486261, 18108370, 18108374, 35696423, 18108385, 22279000
2726	86880588 (5451, 5452)	Novel Protein sim. GBank gl 3342738 (AC005328) - R28860_1, partial CDS [Homo sapiens]		MHC	264486, 264828, 264685

2727	91010470 (5453, 5454)	Novel Protein sim. GBank g1731287 sp P39219 PLUA_ECOLI - RIBOSOMAL LARGE SUBUNIT PSEUDOURIDINE SYNTHASE A (PSEUDOURIDYLATE SYNTHASE) (URACIL HYDROLYASE)	Contains protein domain (PF00048) - RNA pseudouridylylase synthase	synthase	65274572, 5618257, 22278994, 56994075, 22278998, 22278997, 22278999, 80432049, 29331822, 29331824, 66714117, 29331828, 29331827, 35896052, 29331828, 33856870, 264509, 66712502, 284910, 33657402, 60433438, 264758, 55812038, 21906754, 33857084, 55811388, 285018, 285019, 264787, 21906765, 21906787, 21906769, 55811957, 35895917, 52644150, 33857073, 33657109, 33657182, 27486281, 27486282, 27486285, 33657349, 55811576, 35896423, 35895855, 284830, 60431850, 284638, 58182323, 87188518, 80432113, 22279000, 284584, 264565
2728	94128022 (5455, 5456)	Novel Protein sim. GBank g1388043 emb CAA91399 - (Z68521) similar to mitochondrial RNA splicing MSR4 like protein; cDNA EST EMBL: C09217 comes from this gene [Caenorhabditis elegans]	Contains protein domain (PF00153) - Mitochondrial carrier proteins	UNCLASSIFIED	18108394, 56182435, 21906767, 55811957, 35895855, 265021, 284690, 284556, 284259, 264557, 29331822, 284559, 284448, 284288
2728	94126024 (5457, 5458)	Novel Protein sim. GBank g12408085 emb CA816300 - (Z89168) putative RNA splicing protein [Schizosaccharomyces pombe]	Contains protein domain (PF00153) - Mitochondrial carrier proteins	UNCLASSIFIED	85274572, 284259, 60432289, 66712502, 56182435, 284448, 284288, 284369, 55811957, 285021, 284557, 60432113
2730	94126026 (5459, 5460)	Novel Protein sim. GBank g1388043 emb CAA91399 - (Z68521) similar to mitochondrial RNA splicing MSR4 like protein; cDNA EST EMBL: C09217 comes from this gene [Caenorhabditis elegans]	Contains protein domain (PF00153) - Mitochondrial carrier proteins	transport	284087, 29331824, 29331828, 35896052, 264107, 58182435, 285008, 285009, 284592, 80431735, 285011, 284601, 285017, 18108351, 284288, 28148627, 55811957, 265021, 264690, 18108388, 18108374, 264557, 284558, 284559, 18108387, 56526486, 284556, 284488
2731	87723022 (5461, 5462)	Novel Protein sim. GBank g1172339 p C10168 YAU8_SCHPO - HYPOTHETICAL 35.7 KD PROTEIN C28A3.11 IN CHROMOSOME I	Contains protein domain (PF00153) - Carbon-nitrogen hydrolase		264259, 35896052, 285008, 284758, 284762, 284448, 284288, 28148627, 21908769, 87188518, 22279002
2732	94126028 (5463, 5464)	Novel Protein sim. GBank g1388043 emb CAA91399 - (Z68521) similar to mitochondrial RNA splicing MSR4 like protein; cDNA EST EMBL: C09217 comes from this gene [Caenorhabditis elegans]	Contains protein domain (PF00153) - Mitochondrial carrier proteins	UNCLASSIFIED	284087, 284488, 18108358, 58182435, 284689, 35896423, 55811957, 265021, 285008, 285009, 284910, 285009, 284690, 284555, 284259, 284556, 264657, 284558, 284559, 18108383, 33857109, 87168518, 265010, 284601, 60432113, 285017, 284905, 284448, 263972, 284369, 284567
2733	87383080 (5465, 5468)	Novel Protein sim. GBank g1451982 p b BAA75970.1 - (AB017814) OASIS protein [Mus musculus]	Contains protein domain (PF00170) - bZIP transcription factor	UNCLASSIFIED	29331825, 284509, 284809
2734	94140288 (5467, 5468)	Novel Protein sim. GBank g1451982 p b BAA75970.1 - (AB017814) OASIS protein [Mus musculus]	Contains protein domain (PF00170) - bZIP transcription factor	transcription factor	60424179, 52644507, 56182575, 284259, 29331828, 284907, 284510, 284910, 60433358, 265019, 55811576, 284881, 284783, 284687, 33857182, 18108370, 60431528, 80431850, 56182323, 83373044

2735	67712336 (5488, 5470)	Novel Protein sim. GBank gij3650568 (AC005276) - ESTs gbT721276, gbT745403, and gbAA586113 come from this gene. [Arabidopsis thaliana]		glycoprotein	22276886, 60432289, 29331827, 29146498, 264108, 264909, 264112, 33657402, 67186474, 265017, 264762, 264448, 264704, 264664, 21906765, 264693, 33657109, 263976, 264638, 264638, 264557, 22278000, 22278002, 264567
2736	80247855 (5471, 5472)			UNCLASSIFIED	264905, 264628, 264629, 263978, 264632, 264564
2737	87604526 (5473, 5474)	Novel Protein sim. GBank gij2556501[db]BA422896] - (O53850) hepatoma-derived growth factor [Mus musculus]			264680
2738	85731806 (5475, 5476)	Novel Protein sim. GBank gij5420367[emb]CAB46679.1] - (A2343459) proteolipophogyan [Leishmania major]		UNCLASSIFIED	264468, 265008, 264768, 264691
2740	94148762 (5478, 5480)	Novel Protein sim. GBank gij3471386[emb]CAA75495] - (Y15197) microtubule-associated protein, MAP-115 [Mus musculus]		UNCLASSIFIED	264684, 63373044, 264566
2741	88047516 (5481, 5482)	Novel Protein sim. GBank gij3242764 (AC005154) - similar to protein UZ6928 (P1D: g861308) [Homo sapiens]		UNCLASSIFIED	264468, 56162575, 22276995, 35696286, 22276997, 22276998, 22276999, 264259, 29331822, 29331824, 29331825, 29331827, 35696052, 29331826, 29146498, 29331830, 265006, 265007, 265008, 60432229, 33657402, 55612036, 67186474, 265010, 265011, 265017, 265018, 265019, 264805, 264681, 264288, 264389, 52844229, 21906765, 21906766, 21906767, 21906768, 21906769, 265020, 265022, 264691, 264692, 33657109, 16106370, 16106374, 55810784, 35695855, 264634, 60431650, 264639, 56162323, 16106362, 16106365, 65274727, 22279002, 264564
2742	87646844 (5483, 5484)	Novel Protein sim. GBank gij4758412[ref]NP_004472.1]pGALN - UDP-N-acetyl-alpha-D-galactosamine; polypeptide N-acetylglucosaminyltransferase 2 (GALNAc-T2)	Contains protein domain (PF00652) - Similarity to lectin domain of ricin beta-chain, 3 copies.	transferase	22276996, 52644045, 52644229, 21906768, 21906769, 265020, 60170615, 264691
2743	87627891 (5485, 5486)	Novel Protein sim. GBank gij468311[emb]CAB37892] - (AL031432) d4463N24.2.1 (PUTATIVE novel protein) (isoform 1) [Homo sapiens]		UNCLASSIFIED	264259, 264805, 264758, 55612038, 264369, 29146627
2744	94126030 (5487, 5488)	Novel Protein sim. GBank gij3860433[emb]CAA91399] - (Z68521) similar to mitochondrial RNA splicing MSR4 like protein, cDNA EST EMBL:CO9217 comes from this gene [Caenorhabditis elegans]	Contains protein domain (PF00153) - Mitochondrial carrier proteins	kinase	35696286, 264259, 264906, 264908, 265006, 60433436, 265017, 16106331, 264446, 264764, 264288, 21906765, 21906767, 264690, 264691, 264693, 263989, 263971, 35695855, 264637, 264556, 16106362, 60432113

2745	87740123 (5488, 5490)	Novel Protein sim. GBank gll4403759jg/AAD189261 - (AF036863) RNA helicase [Homo sapiens]	Contains protein domain (PF00271) - Helicase domain	UNCLASSIFIED	35686236, 264508, 264905, 264907, 264908, 264909, 264510, 264512, 265008, 264758, 264801, 265017, 264804, 264783, 264288, 264888, 264789, 264893, 35696423, 35695855, 264634, 264636, 264563, 264564, 264565
2746	85418801 (5487, 5492)	Novel Protein sim. GBank gll4758728j/efjNP_004680.1/pMTA1 - metastasis associated 1	Contains protein domain (PF00320) - GATA zinc finger	UNCLASSIFIED	22276896, 22276998, 22278893, 29331822, 29331826, 29331827, 35696052, 29331828, 264905, 264906, 264907, 264908, 264909, 52644045, 265006, 60170831, 264596, 55812038, 265018, 264883, 264268, 21806785, 21906767, 21906768, 21906769, 265020, 264890, 33857023, 264693, 33857109, 18108368, 18108374, 264558, 18108385, 22279000, 284563
2747	84112677 (5483, 5494)	Novel Protein sim. GBank gll457803j/efjNP_000262.1/pNP-C1 - Niemann-Pick disease, type C1		glycoprotein	264569, 52644507, 18108394, 22278955, 35696286, 22276897, 22278998, 52645080, 29331824, 56162161, 29331828, 29331827, 35696052, 264907, 264908, 264909, 265009, 33108954, 55811388, 87168474, 265010, 87168559, 264603, 265018, 264760, 264686, 264768, 21906768, 35695917, 60170815, 264692, 33657023, 52645128, 27488284, 60431528, 18108374, 35696423, 35695855, 264558, 56162323, 18108385, 264482
2748	81214983 (5485, 5486)	Novel Protein sim. GBank gll4181272j/emb/CAA099841 - (AJ012295) apag protein [Rizobium etli]	Contains protein domain (PF00846) - F-box domain.		65274572, 28331828, 284112, 284511, 265018, 264760, 264767, 264768, 264769, 21806768, 21906769, 265020, 27488282, 56526488, 87168516, 22279000
2749	87346307 (5487, 5488)				264258, 264908, 264510, 265008, 265009, 264760, 264368, 264768, 264563
2750	87336344 (5489, 5500)	Novel Protein sim. GBank gll1872498 (U74287) - PLUS [Oryctolagus cuniculus]		UNCLASSIFIED	264488, 52644507, 18108386, 58994075, 264259, 29331825, 29331828, 29331827, 29331828, 264508, 265009, 264910, 264591, 264595, 33657084, 265011, 265019, 16108351, 284288, 284686, 284768, 264689, 55811957, 264693, 27488284, 18108370, 18108374, 264558, 18108385, 264482, 264563
2751	87057465 (5501, 5502)			UNCLASSIFIED	29331822, 29331824, 285017, 33657023
2752	88082679 (5503, 5504)	Novel Protein sim. GBank gll3041859 (AC004534) - CG-2 homeodomain protein-like; similar to UES067 (P1Dg1575528) [Homo sapiens]	Contains protein domain (PF00046) - Homeobox domain	homeobox	

2753	84138872 (5505, 5506)	Novel Protein sim. GBank gl3651648 (AF098301) - neural F-box protein NF942 [Rattus norvegicus]	Contains protein domain (PF00646) - F-box domain.	UNCLASSIFIED	56182575, 56994075, 22278898, 29331822, 29331824, 29331825, 29331826, 265007, 264593, 56812038, 33109854, 18108351, 264288, 56181562, 21906787, 21908788, 265021, 264693, 18108374, 65274781, 264632, 56182323, 22279002, 264583, 264587
2754	84115513 (5507, 5508)	Novel Protein sim. GBank gl355428 (U13738) - calmodulin-like protein [Pisum sativum]	Contains protein domain (PF00036) - EF hand	struct	22278898, 88714117, 29331827, 35696052, 29331828, 264908, 264908, 264908, 265011, 265017, 265018, 265019, 264288, 21908785, 21906787, 265022, 33657023, 264693, 56182323, 18108382, 22279000
2755	88001472 (5509, 5510)	Novel Protein sim. GBank gl2988853 (AC004510) - R30385_2 [Homo sapiens]	Contains protein domain (PF00096) - Zinc finger, C2H2 type	transcript/leader	264594
2756	11465908 (5511, 5512)	Novel Protein sim. GBank gl1173539 (U30473) - putative src-like adapter protein; non-catalytic src-like adapter protein containing SH3 and SH2 domains, homolog of mouse SLAP. Method: conceptual translation supplied by author [Homo sapiens]	Contains protein domain (PF00017) - Src homology domain 2	UNCLASSIFIED	85658542, 264693
2757	85361590 (5513, 5514)	Novel Protein sim. GBank gl3608372 (AF053768) - brain specific cortactin-binding protein CBP90 [Rattus norvegicus]	struct	264693	65274572, 35696288, 88714117, 29331828, 264506, 56182435, 21906754, 56811857, 264628, 264638, 56182323, 22279002
2758	91005312 (5517, 5518)	Novel Protein sim. GBank gl2072200 (U94863) - p40 [Borna disease virus]	Contains protein domain (PF00023) - Anky repeat	UNCLASSIFIED	22278898, 28331822, 28331830, 265010, 265019, 264288, 21906785, 21908786, 21906788, 265020, 56182323, 22279002, 264583
2759	78824796 (5519, 5520)	Novel Protein sim. GBank gl4814573 (Jeml)CAB43685.1 - (AL050380) hypothetical protein [Homo sapiens]	Contains protein domain (PF00889) - PAS domain	struct	18108394, 22278898, 264906, 264908, 265006, 265007, 264757, 265010, 265011, 285017, 265019, 18108351, 284446, 264683, 284888, 264768, 265020, 265021, 265022, 264691, 18108362, 264693, 18108385, 33657108, 18108386, 16106370, 16108381, 18108382, 18108384, 18106386, 67186516, 264368, 35688423
2760	87582889 (5523, 5524)	Novel Protein sim. GBank gl3136150 (AF050182) - PERIOD 3 [Mus musculus]	Contains protein domain (PF01352) - KRAB box	transcript/leader	52646642, 264258, 29331822, 29331825, 29331826, 28331826, 33656870, 264905, 264907, 29331630, 285006, 265009, 21906754, 285019, 264448, 21908789, 27488282, 56182323, 56528466, 87188518, 264487
2761	87582889 (5523, 5524)	Novel Protein sim. GBank gl3136150 (AF050182) - PERIOD 3 [Mus musculus]	Contains protein domain (PF01352) - KRAB box	transcript/leader	52646642, 264258, 29331822, 29331825, 29331826, 28331826, 33656870, 264905, 264907, 29331630, 285006, 265009, 21906754, 285019, 264448, 21908789, 27488282, 56182323, 56528466, 87188518, 264487
2762	87582889 (5523, 5524)	Novel Protein sim. GBank gl3136150 (AF050182) - PERIOD 3 [Mus musculus]	Contains protein domain (PF01352) - KRAB box	transcript/leader	52646642, 264258, 29331822, 29331825, 29331826, 28331826, 33656870, 264905, 264907, 29331630, 285006, 265009, 21906754, 285019, 264448, 21908789, 27488282, 56182323, 56528466, 87188518, 264487
2763	67538986 (5525, 5526)	Novel Protein sim. GBank gl3511122 (AF060503) - zinc finger protein [Homo sapiens]	Contains protein domain (PF00037) - FGCGY family of carbohydrate kinases	transcript/leader	52646642, 264258, 29331822, 29331825, 29331826, 28331826, 33656870, 264905, 264907, 29331630, 285006, 265009, 21906754, 285019, 264448, 21908789, 27488282, 56182323, 56528466, 87188518, 264487
2764	84305140 (5527, 5528)	Novel Protein sim. GBank gl2905643 (AF045244) - ribitol kinase [Klebsiella pneumoniae]	Contains protein domain (PF00037) - FGCGY family of carbohydrate kinases	transcript/leader	52646642, 264258, 29331822, 29331825, 29331826, 28331826, 33656870, 264905, 264907, 29331630, 285006, 265009, 21906754, 285019, 264448, 21908789, 27488282, 56182323, 56528466, 87188518, 264487

2765	94315105 (5529, 5530)	Novel Protein sim. GBank gll4686672[emb]CA17888.2 - (ALU22016) /prediction=(method:: /prediction=(method:: /match=(desc: [Drosophila melanogaster])			264488, 65274572, 22278895, 35686288, 22278898, 22278899, 264259, 29331822, 28331824, 35686052, 264506, 264509, 284905, 264806, 264807, 264808, 264809, 264512, 264910, 265008, 264592, 264595, 264758, 55812036, 33109854, 265010, 87168558, 264600, 265016, 264780, 264781, 264762, 264763, 264448, 264764, 264268, 264788, 264767, 264788, 56181582, 21908764, 21908765, 21908768, 35685917, 285021, 264891, 264892, 33657023, 33657109, 264626, 16108370, 264629, 18108374, 55811576, 35686423, 35685855, 264631, 264634, 264635, 264836, 264837, 264638, 264538, 264839, 63373044, 87168518, 22279000, 22279002, 264583, 264482
2766	94315109 (5531, 5532)	Novel Protein sim. GBank gll541611[emb]CAB46854.1 - (A38655) hypothetical protein [Cans familiaris]	Contains protein domain (PF00076) - RNA recognition motif, (a.k.a. RRM, RBD, or RNP domain)	dna_ma_bind	264488, 65274572, 60432289, 264807, 264908, 264511, 264512, 60433356, 264286, 264685, 264689, 35685917, 285022, 264693, 264828, 65274781, 264635, 264555, 264556, 264557, 264636, 264558, 264559, 63373044, 60432113
2767	80204287 (5533, 5534)	Novel Protein sim. GBank gll1078451[emb]A55463 - tropomodulin, skeletal muscle - chicken		struct	264112, 263874, 264538
2766	94322238 (5535, 5536)	Novel Protein sim. GBank gll541322[emb]CAB46721.1 - (AL031427) GJ167A19.1 (novel protein) [Homo sapiens]		UNCLASSIFIED	264486, 283994, 264488, 35686288, 22278898, 22278899, 264259, 28331822, 60432289, 28331828, 35686052, 264508, 264509, 264805, 264806, 264807, 28331830, 264808, 264809, 264510, 264511, 264512, 265008, 264810, 265009, 264581, 264582, 264593, 33657402, 264594, 264595, 264757, 264598, 285011, 264604, 265018, 264760, 264881, 18108351, 264764, 264268, 264786, 264688, 264786, 16108357, 264789, 264889, 21908786, 21908788, 21908789, 35685917, 265021, 265022, 33657023, 33657109, 264828, 264629, 16108374, 35686423, 35685855, 264630, 264831, 264632, 264634, 264835, 264636, 264555, 264638, 264639, 63373044, 264563, 264564, 264565, 264566, 264468, 264587

2769	95311088 (5537, 5539)	Novel Protein sim. GBank glij41985jtemb[CAB46375.1 - (AL086725) hypothetical protein (Homo sapiens)]		tubulin	264488, 56182575, 264259, 29331822, 29331824, 68714117, 29331825, 60432289, 29331828, 264908, 29331830, 56182435, 264592, 33657402, 264448, 264369, 264288, 60170615, 264691, 33657023, 264692, 33657109, 18108374, 55811576, 264634, 264836, 56182323, 83073044, 80432113, 18108398, 22278995, 22278998, 22278999, 264105, 265006, 285019, 18108351, 264687, 21906785, 18108364, 264629, 18108374, 264631, 18108385, 18108388, 264259, 28331822, 29331824, 29331825, 264369, 284259, 28331822, 29331824, 68714117, 60432289, 29331827, 264905, 265009, 284592, 55812038, 85274444, 264766, 21906769, 33657109, 263978, 264555, 264638, 264557, 83373044, 264583, 264564, 264488, 18108374, 264686, 264687, 263978, 56182435, 264689, 55810784, 21908768, 35896423, 55811576, 55274791, 56181886, 55811957, 35895855, 264110, 265021, 264112, 265022, 265008, 265008, 264092, 284094, 60431850, 264637, 264691, 264259, 264692, 263981, 264594, 80433356, 56182323, 264693, 264757, 58182181, 55812038, 29331825, 80424269, 18108385, 29331828, 29331827, 27486281, 29331828, 35898052, 55811386, 264107, 60432113, 265017, 55811150, 18108351, 264681, 264908, 18108370, 264484, 264882, 20281069, 264448, 88712502, 264693, 264784, 264288, 264684, 264768, 283974, 22278995, 35896286, 22278998, 22278999, 264259, 29331828, 80432289, 35896052, 264112, 33657402, 21908754, 87168559, 265017, 265018, 264288, 21906768, 21906787, 21906768, 35895917, 285020, 285021, 33657023, 33657109, 18108370, 263978, 35896423, 35895855, 87188518, 22278000, 264482
2770	87730182 (5539, 5540)	Novel Protein sim. GBank glij5701963jtemb[CAB52157.1 - (AL109738) WD repeat protein (Schistosoma mansoni)]	Contains protein domain (PF00400) - WD domain, G-beta repeat	kinase	
2771	88094071 (5541, 5542)	Novel Protein sim. GBank glij3083433 (AC004125) - Unknown gene product (Homo sapiens)		UNCLASSIFIED	
2772	95357308 (5543, 5544)	Novel Protein sim. GBank glij488553jtemb[NP_005465.1]pnt C - histone deacetylase 5	Contains protein domain (PF00850) - histone Histone deacetylase family		
2773	84138984 (5545, 5546)	Novel Protein sim. GBank glij288888 (AC005253) - R26445_1 (Homo sapiens)		UNCLASSIFIED	
2774	87619908 (5547, 5548)	Novel Protein sim. GBank glij485852jtemb[P24389]LTS3_CAEL - HYPOTHETICAL TO.7 KD PROTEIN F09G.3 IN CHROMOSOME III	Contains protein domain (PF00380) - Ribosomal protein S9/S18	ribosomalpro	

2775	95307987 (5549, 5550)	Novel Protein sim. GBank gik4689132ipbIAD2775, 1(AF07704 - (AF077042) 30S ribosomal protein S7 homolog [Homo sapiens]	Contains protein domain (PF00177) - ribosomal prot	264468, 22278995, 56994075, 22276986, 35696266, 22276998, 22278999, 264259, 29331822, 29331824, 29331825, 80432289, 29331827, 29331828, 35698052, 265007, 21906754, 285017, 265019, 264448, 264682, 264369, 264286, 18108354, 52644229, 264689, 21906765, 21906766, 21906767, 21906768, 21906769, 35695917, 265021, 265022, 60170615, 264691, 18108370, 35696423, 65274781, 35695855, 264634, 60431850, 60170394, 56162323, 264558, 18108388, 22279000, 264563, 264565, 264468, 264567
2776	87791557 (5551, 5552)			58182575, 22278998, 22278999, 264259, 29331822, 29331824, 284908, 29331830, 264510, 33657402, 21906754, 55811386, 285017, 285019, 284448, 264286, 21906765, 21906766, 21906767, 21906768, 21906769, 265020, 285021, 285022, 60170615, 55810764, 55811576, 264555, 56526486, 22279000
2777	79618729 (5553, 5554)		UNCLASSIFIED	264907, 264766
2778	82112411 (5555, 5556)	Novel Protein sim. GBank	UNCLASSIFIED	264907, 264593, 264760, 264626
2779	87649729 (5557, 5558)	Novel Protein sim. GBank gik4680711gipbIAD27745, 1(AF13297 - (AF132970) CGI-36 protein [Homo sapiens]	UNCLASSIFIED	22278997, 264259, 29331824, 86714117, 35698052, 265006, 264512, 264448, 264286, 29148627, 18108364, 20281149, 18108370, 264629
2780	94678397 (5559, 5560)	Novel Protein sim. GBank gik4758524ipbIAD27745, 1(AF13297 - (AF132970) CGI-36 protein [Homo sapiens]	Contains protein domain (PF00760) - kinase	29146499, 65274781, 264634, 264639
2781	91220057 (5561, 5562)	Novel Protein sim. GBank gik4689352gipbIAD212221 - (AF068502) ubiquitin specific protease UBP-43 [Mus musculus]	CNH domain Contains protein domain (PF00443) - Ubiquitin carboxy-terminal hydrolase family 2	60424178, 29331824, 60424269, 66714117, 29331826, 56182435, 87188474, 265017, 264784, 56181582, 21906765, 21906766, 21906768, 35695917, 265020, 33657023, 35695855, 56182323, 87188518
2782	94233146 (5563, 5564)	Novel Protein sim. GBank gik4505073ipbIAD27745, 1(AF13297 - (AF132970) CGI-36 protein [Homo sapiens]	Contains protein domain (PF00560) - Leucine Rich Repeat	65274572, 22278998, 22278999, 80432049, 264259, 29331822, 29331824, 29331826, 80432289, 29331826, 264805, 264807, 264906, 264909, 52644045, 265009, 265017, 265018, 264804, 265019, 264760, 264683, 264286, 264768, 264685, 264686, 264766, 52644229, 264689, 21906766, 265020, 265021, 264691, 18108362, 264692, 33657023, 264693, 33657109, 33657349, 16108370, 264628, 263976, 35695855, 264557, 56162323, 80373044, 18108365

2783	80016629 (5565, 5568)	Novel Protein sim. GBank g 1728831 sp P39188 ALU1_HUMAN - III ALU SUBFAMILY J WARNING ENTRY III		tm7	264809, 264628, 283976, 263981
2784	87614380 (5567, 5568)				264259, 28331822, 28331824, 28331825, 264482
2785	88071830 (5568, 5570)	Novel Protein sim. GBank g 12134833 p h S58890 - collapse response mediator protein - human		UNCLASSIFIED	22278899, 264608, 264758, 265018, 264769, 21906795, 21908768, 21908769, 265020, 264584
2786	87406542 (5571, 5572)	Novel Protein sim. GBank g 2073564 (U80223) - eukaryotic initiation factor eIF-2 alpha kinase: DGCM2 (Drosophila melanogaster)	Contains protein domain (PF00069) - Eukaryotic protein kinase domain	Kinase	264905, 265017
2787	87901288 (5573, 5574)	Novel Protein sim. GBank g 15174507 p pNP_008020.1 pMA11 - paraneoplastic neuronal antigen		UNCLASSIFIED	264768, 21908765, 21908768, 55811957, 22278899, 264093, 60170615, 264259, 29331822, 18108365, 28331824, 33857108, 29331827, 35698052, 284100, 264105, 264808, 263977, 55811576, 264635, 264637, 80433438, 80432113, 285017, 285019, 22279002, 55811150, 284369, 284268
2788	88080844 (5575, 5576)	Novel Protein sim. GBank g 3252828 (AC004382) - Unknown gene product [Homo sapiens]		UNCLASSIFIED	264103, 21906769, 264693
2789	85491275 (5577, 5578)	Novel Protein sim. GBank g 12485729 p Q92558 Y281_HUMAN - HYPOTHETICAL PROTEIN KIAA0281 (HAB725)		UNCLASSIFIED	264468, 264769, 264810, 264831, 284636, 264680, 264691, 264258, 264638, 28331824, 264508, 264509, 264905, 284563, 284908, 284628, 18108370, 284907, 264784, 264808, 264288, 264909
2790	87602784 (5578, 5580)	Novel Protein sim. GBank g 15101772 emb CAB45135.1 - (A242978) p821 [Homo sapiens]		struct	21906784, 18108368
2791	88083185 (5581, 5582)	Novel Protein sim. GBank g 2911288 (AC002550) - Unknown gene product [Homo sapiens]			22278898, 22278897, 264259, 28331822, 28331824, 28331825, 66714117, 28331826, 60432289, 28331827, 35698052, 284906, 68712502, 28331830, 264908, 80432228, 60433356, 80433438, 33108954, 285010, 265011, 265017, 265016, 265019, 284682, 284448, 264288, 264369, 264766, 52844228, 21908785, 21908786, 285020, 285021, 33657023, 263974, 18108374, 65274781, 35895655, 284838, 284558, 264558, 58182323, 83373044, 18108385, 58526486, 60432113, 22279000, 264587
2792	85083783 (5583, 5584)	Novel Protein sim. GBank g 12854163 p AA002581.1 - (AF045642) No definition line found [Caenorhabditis elegans]			264259, 60432288, 66712502, 265009, 264636
2793	87425476 (5585, 5588)			UNCLASSIFIED	264259, 60432288, 66712502, 265009, 264636
2794	85794830 (5587, 5588)			UNCLASSIFIED	284888, 285022, 264691, 18108388, 284567

2795	95334888 (5589, 5590)	Novel Protein sim. GBank gij454146 refNP_006346.1 PUBEE2 - ubiquitin-conjugating enzyme E2E 3 (homologous to yeast UBC45)	Contains protein domain (PF00179) - Ubiquitin Ubiquitin-conjugating enzyme	Ubiquitin	65274572, 56182575, 35696286, 22276996, 22278999, 60432049, 264259, 28331822, 28331823, 66714117, 60432289, 28331826, 28331827, 35696052, 28331828, 204907, 66712502, 56182435, 264511, 285007, 264512, 264910, 60432229, 60433356, 60433438, 33108954, 85858542, 265016, 265019, 264288, 264688, 21906784, 21906785, 21906786, 21906787, 21906789, 55811857, 265020, 265021, 285022, 52844150, 33657023, 264683, 65274620, 33657109, 35696423, 55811578, 85274791, 56182323, 56526486, 60432113, 22279002, 264482, 264563, 264484, 264567
2786	94846857 (5591, 5592)	Novel Protein sim. GBank gij4680851 gb AAD27715.1 AF13294 - (AF132940) CGI:08 protein [Homo sapiens]		UNCLASSIFIED	18108394, 65274572, 56182575, 56894075, 22278999, 264490, 60432049, 264259, 29331822, 29331824, 29331826, 35696052, 264509, 264908, 264907, 264908, 66712502, 56182435, 264510, 285008, 264512, 285007, 285008, 264810, 285009, 264581, 264582, 60433356, 264594, 264585, 55812036, 264598, 21906754, 60174639, 87168474, 265010, 285011, 285017, 285018, 285019, 55611150, 264782, 18108351, 264448, 264682, 264368, 264288, 264685, 264768, 264687, 56181562, 264769, 264689, 21906765, 21906768, 21906767, 21906768, 21906769, 55811957, 35695917, 265020, 265021, 60170615, 52844150, 264692, 33657023, 18108362, 264693, 65274620, 33657109, 33657182, 27468265, 33657349, 18108374, 35696423, 65274791, 35695855, 264558, 264557, 56182323, 264558, 80170394, 83373044, 65274727, 67168518, 22279000, 264563, 264584, 264565, 264585, 264587
2797	95110790 (5593, 5594)	Novel Protein sim. GBank gij4638557 gb AAD31040.11 - (AF143859) potassium channel modulatory factor DEBT-91 [Mus musculus]	Contains protein domain (PF00569) - Zinc finger present in dystrophin, CBP/p300	UNCLASSIFIED	56182575, 22278995, 22278996, 22278997, 28331827, 28146489, 264509, 264908, 56182435, 264757, 21906754, 265010, 265017, 265019, 264681, 264682, 264683, 264686, 21906765, 21906767, 21906768, 21906769, 28146628, 265020, 52844150, 264690, 33657182, 264629, 18108376, 56182323, 22279002, 264563
2788	86198005 (5595, 5596)	Novel Protein sim. GBank gij2652845 (AF007160) - unknown [Homo sapiens]		UNCLASSIFIED	264486, 264480, 264259, 264448, 20281149, 20281152, 264558, 264557, 264558, 264559, 264463, 264468, 264567

2789	88080651 (5587, 5598)	Novel Protein sim. GBank glij3255625 (AC004382) - Unknown gene product [Homo sapiens]		UNCLASSIFIED	22278996, 22278997, 22278998, 22278999, 60432048, 264259, 29331826, 35698052, 29331826, 264809, 60433358, 33657402, 33108954, 87188474, 264446, 52644229, 21806768, 21906787, 21906768, 35695917, 265020, 265021, 265022, 52644150, 35695855, 264634, 60432113, 22279000
2800	66318481 (5589, 5600)	Novel Protein sim. GBank glij4240301[id]BAA74929.1 - (AB02073) KIAA0808 protein [Homo sapiens]		glycoprotein	264468, 56894075, 264259, 20281089, 29331825, 29331827, 264905, 56182435, 265006, 265011, 87186559, 265017, 265019, 264446, 264288, 264766, 264686, 60170615, 264691, 264692, 27486265, 264826, 264629, 264636, 264557, 264558, 264559, 67188518, 264584, 264588, 264587
2801	86068614 (5601, 5602)		Contains protein domain (PF00827) - UBA domain		265007, 264687
2802	88082477 (5603, 5604)	Novel Protein sim. GBank glij237865 (AC002484) - organic cation transporter, 50% similarity to JC4864 (P1D:g2143892) [Homo sapiens]	Contains protein domain (PF00083) - Sugar (and other) transporter	transport	264446, 35695855
2803	79577446 (5605, 5606)	Novel Protein sim. GBank		UNCLASSIFIED	264839
2804	57111131 (5607, 5608)	glij4558368[id]AA023029.1 (AC006565) hypothetical protein [Arabidopsis thaliana]	Contains protein domain (PF01585) - G-patch domain	peptidase	264566
2805	67398466 (5609, 5610)				264092, 264259, 29331822, 29331824, 29331827, 29331826, 264508, 285007, 265009, 21906754, 264369, 264288, 264632, 60170394, 264583, 264482
2806	87888951 (5611, 5612)	Novel Protein sim. GBank glij1168873[id]P44403[CLPB_HAEN - CLPB PROTEIN		UNCLASSIFIED	22278995, 22278996, 22278998, 264259, 29331822, 29331824, 29331828, 29331827, 29146498, 265008, 265009, 60433436, 265017, 265018, 265019, 264446, 264288, 21906765, 21906787, 21906769, 29146628, 265022, 52644150, 56182323

2807	B1720702 (5813, 5614)	Novel Protein sim. GBank g14466310genb(CA33791) - (AL031432) d1485N24.1 (PUTATIVE novel protein similar to predicted yeast and worm proteins) [Homo sapiens]		UNCLASSIFIED	52644507, 52645158, 52646842, 18108398, 56182575, 22278994, 22278995, 22278998, 56994075, 35698286, 22278997, 22278998, 22278999, 264259, 52845080, 29331822, 29331824, 29331825, 29331828, 35898052, 29331828, 33856970, 264100, 264105, 264907, 52644045, 60433358, 264594, 60433436, 52646317, 21908754, 33109954, 33657084, 52844298, 87188474, 285011, 87188559, 265017, 265018, 265019, 18108351, 264448, 264783, 264887, 52644229, 21906765, 21908766, 21906767, 21908768, 21908769, 35695917, 265021, 265022, 52644150, 33657023, 52645129, 33657109, 33657182, 27486281, 27486282, 27486284, 27486285, 35895783, 18108370, 18108372, 18108374, 18108376, 283977, 18108377, 35698423, 35895855, 52644332, 83373044, 18108385, 18108387, 87188518, 60432113
2808	95359111 (5615, 5618)	Novel Protein sim. GBank g15541863jenb(CA35107.1) - (AL096857) hypothetical protein [Homo sapiens]		MHC	60432289, 284510, 285010, 265017, 265018, 264681, 264686, 265021, 284890, 22279000, 264568
2809	88083530 (5617, 5618)	Novel Protein sim. GBank g1272581 (AC004002) - similar to ciliary dynein beta heavy chain. 76% Similarity to P23098 (P1Dg118865) [Homo sapiens]		ATPase-associated	18108351
2810	87259032 (5618, 5620)			UNCLASSIFIED	264569, 22278996, 264091, 264259, 29331828, 29146499, 29146629, 29148784, 264693
2811	91235845 (5621, 5622)			UNCLASSIFIED	22278999, 264259, 66712502, 264693
2812	88093334 (5623, 5624)	Novel Protein sim. GBank g13284583 (AC005189) - match to ESTs H97758 (NIDg118843) and AA085546 (NIDg1828773) [Homo sapiens]		UNCLASSIFIED	264106
2813	91218735 (5825, 5628)	Novel Protein sim. GBank g14240273jdbj(AA74915.1) - (AB020889) KIAA0892 protein [Homo sapiens]		UNCLASSIFIED	56182575, 29331828, 264906, 68712502, 55811388, 285017, 285018, 265019, 264683, 265020, 87168518, 60432113
2814	90880908 (5827, 5628)	Novel Protein sim. GBank g13546781 (AC005620) - R33590.1 [Homo sapiens]	Contains protein domain (PF00170) - bZIP transcription factor	transcription factor	65274572
2815	79774521 (5628, 5630)			UNCLASSIFIED	264907, 264909
2816	85358228 (5631, 5632)	Novel Protein sim. GBank g1542038jenb(CA346880.1) - (AJ243460) proteophosphoglycan [Leishmania major]		UNCLASSIFIED	264468, 35898286, 29331825, 29331828, 264508, 264509, 264905, 264906, 264907, 264908, 264909, 264511, 264910, 265011, 264760, 264881, 264788, 284789, 264688, 21908765, 264693, 264628, 18108370, 264829, 264631, 284834, 264583, 264564, 264565, 264488

2817	87748542 (5633, 5634)	Novel Protein sim. GBank gij1283846 (U56986) - coded for by <i>C. elegans</i> cDNA yk3003.5; coded for by <i>C. elegans</i> cDNA yk3003.3 [<i>Caenorhabditis elegans</i>]				264259, 28331822, 28331827, 264508, 264509, 264905, 264907, 264908, 264909, 56182435, 264510, 265006, 264511, 264512, 264593, 264758, 265010, 264780, 264781, 264784, 264788, 264887, 264789, 55811957, 35693917, 33857109, 263978, 264634, 264638, 264639, 264584, 264585, 264568, 264488, 264567, 86712502
2818	88073579 (5635, 5636)	Novel Protein sim. GBank gij1549986 (U13149) - possible spore-associated protein [Permetium citare]	Contains protein domain (PF00023) - Ank repeat	UNCLASSIFIED		
2819	87793527 (5637, 5638)			UNCLASSIFIED		264909, 264511
2820	87765744 (5639, 5640)	Novel Protein sim. GBank gij49287738b/AAD34147.1/AF15209 (AF15209) CGI-05 protein [Homo sapiens]	Contains protein domain (PF00818) - Uncharacterized protein family UPF0004	UNCLASSIFIED		18108394, 52646365, 52644045, 264112, 265009, 21906754, 265017, 18108351, 264883, 264369, 264888, 264789, 21906769, 33657023, 33657109, 18108370, 18108374, 18108377, 18108385
2821	85320511 (5641, 5642)	Novel Protein sim. GBank gij398144sp/PO2747C1QC_HUMAN - COMPLEMENT C1Q SUBCOMPONENT, C CHAIN PRECURSOR (AB002363) KIAA0385 [Homo sapiens]	Contains protein domain (PF00386) - C1q domain	complement		264638
2822	94260221 (5643, 5644)	Novel Protein sim. GBank gij2224671kb/BA208201 - (AB002363) KIAA0385 [Homo sapiens]		UNCLASSIFIED		264259, 28331822, 60432288, 264805, 60433356, 33857402, 265011, 264780, 21906785, 80170615, 264838
2823	85320513 (5645, 5646)	Novel Protein sim. GBank gij398144sp/PO2747C1QC_HUMAN - COMPLEMENT C1Q SUBCOMPONENT, C CHAIN PRECURSOR	Contains protein domain (PF00386) - C1q domain	complement		264786
2824	95320515 (5647, 5648)	Novel Protein sim. GBank gij398144sp/PO2747C1QC_HUMAN - COMPLEMENT C1Q SUBCOMPONENT, C CHAIN PRECURSOR		UNCLASSIFIED		264907
2825	18742170 (5649, 5650)					264780
2826	94311805 (5651, 5652)	Novel Protein sim. GBank gij385863jemb/CAA220201 - (AL033503) conserved hypothetical protein [Candida albicans]				52645156, 22278894, 35696288, 22278897, 22278998, 52645080, 28331822, 28331824, 28331825, 28331827, 35696052, 28331828, 33856970, 52644045, 52646317, 33657084, 52644286, 265017, 265018, 265019, 264288, 21906764, 21906765, 21906766, 21906767, 21906769, 35695917, 52644150, 33657109, 33657182, 27488261, 27488262, 33657349, 27488265, 35695763, 35696423, 35695855, 52644332, 18108385, 87168516, 264484

2827	85320519 (5653, 5654)	Novel Protein sim. GBank gij399144ipjP02747CICQ, HUMAN - COMPLEMENT C1Q SUBCOMPONENT, C CHAIN PRECURSOR	Contains protein domain (PF00386) - C1q domain	complement	264488, 263994, 264489, 65274572, 29331822, 88714117, 29331827, 29331828, 264508, 264905, 264509, 264906, 264907, 264908, 264909, 264510, 265006, 264511, 285007, 284512, 265008, 264910, 284591, 60432228, 264592, 264593, 264595, 264596, 21908754, 265011, 264600, 264601, 264602, 265017, 264604, 264605, 264288, 264766, 264767, 264689, 55811957, 264534, 264690, 264691, 264692, 264693, 263972, 264629, 35695855, 264631, 264834, 264635, 264555, 264638, 264637, 264556, 264638, 264558, 83373044, 16108385, 60432113, 22278002, 264563, 264564, 264565, 264566, 264587, 264468, 18108391
2828	91229815 (5655, 5656)	Novel Protein sim. GBank gij3588974 (AF077000) - protein tyrosine phosphatase TD14 [Rattus norvegicus]	Contains protein domain (PF00102) - protein tyrosine phosphatase	phosphatase	29331822, 35696052, 264104, 264908, 265007, 264591, 265010, 265011, 265019, 264768, 264688, 55811957, 18108370, 18108374, 55810764, 35696423, 55811576, 56182323, 83373044, 87188518 22278996, 22278997, 264091, 264093, 80432049, 264259, 29331822, 29331825, 29331827, 29331828, 264905, 264509, 66712502, 264510, 264511, 264593, 80433438, 21908754, 265011, 264603, 16106351, 264288, 21908765, 21908768, 21908769, 28148629, 52644150, 264693, 33657109, 16108374, 264834, 18108385, 60432113, 22278000, 264565, 264468 265008, 265019, 264639, 22278002
2829	87651244 (5657, 5658)	Novel Protein sim. GBank gij4880689jgijA027734, iIAF13295 - (AF132959) CGI-25 protein [Homo sapiens]			
2830	88087109 (5659, 5660)	Novel Protein sim. GBank gij2498667jgijO61200INPH1_MOUSE - NEUREXOPHILIN 1			
2831	87614717 (5661, 5662)			UNCLASSIFIED	265017
2832	87631809 (5663, 5664)			UNCLASSIFIED	22278997, 22278999, 52646317, 264288, 264688, 21908767, 80431528, 264638, 22278000
2833	87612938 (5665, 5666)	Novel Protein sim. GBank gij5262615iembjCAB45747.1 - (AL080158) hypothetical protein [Homo sapiens]		UNCLASSIFIED	264555, 264556, 264558
2834	86974703 (5667, 5668)	Novel Protein sim. GBank gij2245587dbjBAC20772 - (AB02311) KIAA0313 [Homo sapiens]			263972
2835	8775712 (5669, 5670)	Novel Protein sim. GBank gij4589532jdjBAA76788.1 - (AB023161) KIAA0844 protein [Homo sapiens]		ATPase associated	60433289, 29331828, 265008, 265010, 265017, 264446, 55811957, 265020, 18108370
2836	85724746 (5671, 5672)	Novel Protein sim. GBank gij2351568 (U76816) - N-RAP [Mus musculus]	Contains protein domain (PF00412) - LIM domain containing proteins	transcriptfactor	264259, 264112, 265010, 264762, 264764, 293974, 264555, 264556, 264559

2637	87766462 (5673, 5674)	Novel Protein sim. GBank gij5420367[emb]CAB46678-11 - (AJ243459) proteophosphoglycan [Leishmania major]		UNCLASSIFIED	18108394, 22278897, 22276996, 29331822, 284907, 284909, 285008, 265007, 265018, 265019, 264682, 264286, 21906766, 21908767, 55811957, 35895917, 16108374, 56182323, 22279000, 22279002
2638	87753392 (5675, 5676)	Novel Protein sim. GBank gij973378 (U31283) - core protein [Hepatitis C virus]		UNCLASSIFIED	16108394, 16108397, 264259, 29331826, 265007, 265019, 264448, 18106366, 20281149, 264565, 284567
2639	85799317 (5677, 5678)	Novel Protein sim. GBank gij1575515 (UB4699) - (thrombospondin-related anonymous protein [Plasmodium gallinaceum])		UNCLASSIFIED	264555
2640	87774665 (5679, 5680)	Novel Protein sim. GBank gij2224605[gb]BAA207801 - (AB002330) KIAA0332 [Homo sapiens]	Contains protein domain (PF00076) - RNA recognition motif, (a.k.a. RRM, RBD, or RNP domain)	dna_ma_bind	264509, 264511, 265011, 264286, 264769, 285020, 264834, 264636, 264556
2641	86982568 (5681, 5682)	Novel Protein sim. GBank gij5578957[emb]CAB51350-11 - (AJ050306) dJ47587.1 (novel KIAA0615 and KIAA0323)		UNCLASSIFIED	264600
2642	80080088 (5683, 5684)	Novel Protein sim. GBank gij5578957[emb]CAB51350-11 - (AJ050306) dJ47587.1 (novel KIAA0615 and KIAA0323)		UNCLASSIFIED	264906, 264907, 264908, 264909, 264910, 264764, 35695855, 63373044, 18108365
2643	81012494 (5685, 5686)	Novel Protein sim. GBank gij5578957[emb]CAB51350-11 - (AJ050306) dJ47587.1 (novel KIAA0615 and KIAA0323)		UNCLASSIFIED	264906, 264907, 264908, 264909, 264910, 264764, 35695855, 63373044, 18108365
2644	56731154 (5687, 5688)	Novel Protein sim. GBank gij5578957[emb]CAB51350-11 - (AJ050306) dJ47587.1 (novel KIAA0615 and KIAA0323)	Contains protein domain (PF00008) - EGF-like domain		284685
2645	94321719 (5689, 5690)	Novel Protein sim. GBank gij5420367[emb]CAB46678-11 - (AJ243459) proteophosphoglycan [Leishmania major]		homeobox	29348498, 87188474, 284886, 35898423, 83373044, 264564
2646	86318613 (5691, 5692)	Novel Protein sim. GBank gij5306263[gb]AAD41995-11(AAC00623) - (AC006233)		UNCLASSIFIED	28331830, 284909, 285008, 265011, 87188659, 264829, 264558
2647	61811757 (5693, 5694)	Novel Protein sim. GBank gij3399676 (AAC005390) - R31180-1 [Homo sapiens]		UNCLASSIFIED	284906, 264766, 264768, 264628, 284837, 284568
2648	67612943 (5695, 5696)	Novel Protein sim. GBank gij5282615[emb]CAB45747-11 - (AL080156) hypothetical protein [Homo sapiens]		UNCLASSIFIED	264490, 264259, 264508, 264905, 264907, 264510, 265007, 285008, 264591, 264592, 284583, 284594, 284595, 55812038, 265010, 265011, 264604, 264763, 264764, 264765, 264786, 264888, 264828, 264829, 284555, 264836, 264556, 264557, 264838, 284558, 264559, 284563, 264586, 264567
2649	86094283 (5697, 5698)	Novel Protein sim. GBank gij3342218 (AC004131) - Unknown gene product [Homo sapiens]		UNCLASSIFIED	29331822, 35698052, 284509, 284906, 265007, 264594, 265018, 264286, 263972, 35698423, 18108394, 58528488, 16106390
2650	87823636 (5699, 5700)			UNCLASSIFIED	35686288, 265008, 265009, 265016, 264286, 35695917, 264893, 18108374, 35695855, 22279000

2851	87820548 (5701, 5702)	Novel Protein sim. GBank gll4321618gblAAD15788.11 - (AF051098) seven transmembrane domain orphan receptor [Mus musculus]		UNCLASSIFIED	264806, 264907, 56182435, 264758, 55611388, 265010, 18108351, 264448, 264369, 21908785, 264691, 264692, 264693, 35695855, 264556, 16108385, 264587, 264591
2852	88967023 (5703, 5704)	Novel Protein sim. GBank gll1825728 (U88308) - similar to drosophila membrane protein PATCHED SP-P18502 (P1Dg128645) [Caenorhabditis elegans]		UNCLASSIFIED	56182575, 55611150, 264690, 27488262, 27486265, 264632, 56182323, 56526486, 22279002
2853	67784830 (5705, 5706)	Novel Protein sim. GBank gll2702347 (AF027503) - putative membrane-associated guanylate kinase 1 [Mus musculus]	Contains protein domain (PF00397) - WW domain	kinase	264806, 264768, 264693, 22279002, 264482
2854	88083557 (5707, 5708)	Novel Protein sim. GBank gll2785825 (AC004021) - kelch protein, ring canal component involved in cytoplasmic bridges; 77% Similarity to A45773 (P1Dg1079096) [Homo sapiens]	Contains protein domain (PF01344) - Kelch motif	dna_ma_bind	35696286, 28331824, 28331828, 28331828, 284908, 264768, 264693, 22279002, 264482
2855	94723856 (5709, 5710)	Novel Protein sim. GBank gll1504040dbjBA4132191 - (D86983) similar to D-melanogaster peroxidasin(U11052) [Homo sapiens]	Contains protein domain (PF00580) - Leucine Rich Repeat	glycoprotein	22278994, 29331822, 29331824, 29331825, 284908, 264908, 265009, 33108954, 265018, 265019, 264448, 21908785, 265020, 264690, 27486265, 83373044, 22279000, 22279002, 264482
2856	88083358 (5711, 5712)	Novel Protein sim. GBank gll3264583 (AC005189) - match to ESTs H87758 (NIDg1118643) and AA085546 (NIDg1628773) [Homo sapiens]			21908766, 22278997, 265022, 28331822, 29331828, 27486282, 285007, 265009, 265017, 264482, 264563, 18108351
2857	95346288 (5713, 5714)	Novel Protein sim. GBank gll3041655 (AC004537) - similar to tumor suppressor p33KNG1; similar to AF044078 (P1Dg2829208) [Homo sapiens]	Contains protein domain (PF00620) - PHD-finger	struct	22278995, 35696288, 29331824, 29331825, 35696052, 264103, 264108, 56182435, 21908765, 21908769, 265020, 18108368, 35695783, 22279002, 264583
2858	87434748 (5715, 5716)	Novel Protein sim. GBank gll462585lpiP35227IME18_HUMAN - DNA-BINDING PROTEIN MEL-18 (ZINC FINGER PROTEIN 144)	Contains protein domain (PF00087) - zinc finger, CCHC type (RING finger)	dna_ma_bind	264568, 264887, 22278995, 22278996, 22278997, 22278999, 264258, 29331826, 29331827, 29331828, 264509, 284805, 264906, 28331630, 284906, 52644045, 284809, 264511, 264512, 265007, 265008, 264910, 265009, 284593, 60433356, 264595, 264758, 21908754, 265010, 265011, 264604, 265018, 264760, 18108351, 264763, 264682, 264764, 264765, 264288, 264369, 264685, 264768, 264768, 18108357, 264769, 21908766, 21908767, 285021, 264534, 60170615, 264691, 264692, 18108370, 264629, 18108374, 264631, 264638, 263981, 18108381, 264558, 18108385, 22279002, 284584, 264586, 264486, 264567
2859	90837675 (5717, 5718)	Novel Protein sim. GBank gll4325320gblAAD17331.11 - (AF124427) claudin-15 [Mus musculus]		UNCLASSIFIED	60424179, 65274572, 28331828, 264905, 264511, 264758, 265011, 21908767, 21908769, 55811957, 265021, 56182323

2860	87532589 (5719, 5720)	Novel Protein sim. GBank gll468186lemh[CAB38414.1] - (AL031588) dJ1163J1.2.1 (novel protein similar to C. elegans B0035.16 end bacterial RNA (5'-Methylaminoethyl-2-thiouridylyle)-Methyltransferases) (isoform 1) [Homo sapiens]		UNCLASSIFIED	264259, 29331822, 29331824, 29331825, 29331826, 29331827, 29331828, 264510, 264511, 33109954, 18108351, 264683, 264765, 264369, 264688, 21906765, 264691, 264892, 264893, 18108388, 22278002, 264482
2881	88898507 (5721, 5722)	Novel Protein sim. GBank gll3941730 (AF108083) - BSA [Homo sapiens]			264369, 264692
2862	87569585 (5723, 5724)	Novel Protein sim. GBank gll4505013lelfp_002310.1pLRLN - leucine-rich neuronal protein		UNCLASSIFIED	264691, 264638
2863	91220421 (5725, 5726)	Novel Protein sim. GBank gll3248088 (AC004473) - Contains similarity to gollath protein gblM87204 from D. melanogaster. [Arabidopsis thaliana]	Contains protein domain (PF00096) - Zinc finger, C2H2 type	struct	56994075, 35686286, 22278998, 29331822, 29331824, 35698052, 29331828, 264108, 264511, 55812038, 33657084, 55811386, 265018, 265019, 21906765, 21906766, 21906769, 35695917, 265020, 265022, 33657023, 33657109, 33657349, 264629, 18108376, 60431850, 56182323, 18108385, 18108387, 87168518, 22279002, 264259, 264910
2884	87420030 (5727, 5728)	Novel Protein sim. GBank gll1073451lpjA55463 - Itopomodulin, skeletal muscle - chicken		struct	22278995, 21906764, 264482
2865	85312181 (5729, 5730)	Novel Protein sim. GBank gll438640 (L19048) - MSA-2 [Plasmodium falciparum]			
2886	95105480 (5731, 5732)	Novel Protein sim. GBank gll565703pplC07086PMP2_RAT - 22 KD PEROXISOMAL MEMBRANE PROTEIN		glycoprotein	85274572, 22278998, 22278999, 22278999, 264259, 29331824, 29331827, 29331828, 60433438, 21906754, 265018, 264488, 264764, 52644150, 83373044
2867	86808001 (5733, 5734)	Novel Protein sim. GBank gll4580987gblAA024571.1pA12108 - (AF121081) cAMP inducible 2 protein [Mus musculus]		UNCLASSIFIED	264488, 264768, 21906768, 22278998, 265022, 264259, 264508, 264905, 264907, 264511, 264910, 264635, 264638, 264637, 265011, 265017, 265018, 265019, 264583, 264088, 264568, 264764, 264389, 264587, 264488, 264288, 264768

2868	95303263 (5735, 5738)	Novel Protein sim. GBank gll1292868 emb CAA63923 - (X94232) i-Cell activation protein [Homo sapiens]			18106392, 264468, 22278994, 22278997, 22278998, 22278999, 60432049, 264259, 29331822, 29331824, 29331825, 60432289, 29331826, 29331827, 35896052, 29331828, 264508, 52644045, 264826, 265006, 265007, 265008, 265009, 264591, 60432229, 284593, 60433356, 264595, 21806754, 265017, 265019, 264682, 264369, 21806765, 21906768, 21906767, 21906768, 285021, 265022, 264681, 33657182, 18106368, 27486261, 27486262, 27486264, 27486265, 18106370, 18106374, 35898423, 35898555, 284832, 58182323, 87188518, 264404, 22279000, 22279002, 284482, 284563, 264564, 264567, 264487
2869	68094412 (5737, 5738)			UNCLASSIFIED	264369
2870	84404574 (5738, 5740)			UNCLASSIFIED	264905, 264908, 264764, 21906769, 264834
2871	68318621 (5741, 5742)	Novel Protein sim. GBank gll5306263 AAD41995.1 AC00623 - (AC00623) unknown protein [Arabidopsis thaliana]		UNCLASSIFIED	264259, 29331822, 60432288, 29331827, 264907, 265008, 265017, 285018, 264682, 264764, 18106354, 285021, 27486285, 264829, 18106387, 264567
2872	85312187 (5743, 5744)	Novel Protein sim. GBank gll112205 p B39066 - proline-rich protein 15 - rat		kinase	263981
2873	88094252 (5745, 5746)			UNCLASSIFIED	264488, 18106374, 264788, 264687, 264888, 264689, 35698423, 35698288, 35695917, 264510, 264511, 265007, 264512, 265008, 264910, 264534, 264834, 264635, 264555, 264592, 264259, 264558, 60433438, 80432289, 35898052, 285011, 264800, 264601, 60432113, 264508, 264563, 264482, 264509, 264905, 264908, 264584, 18106351, 284763, 18106370, 264907, 264588, 284908, 264764, 264286, 264567, 264909, 264486, 284786, 18106391

2874	84313549 (5747, 5748)	Novel Protein sim. GBank gij3212854 (AC004005) - unknown protein [Arabidopsis thaliana]		UNCLASSIFIED	52844507, 52845156, 52846365, 56182575, 22278994, 22278995, 56694075, 35696286, 60432049, 284259, 52845080, 29331822, 56182181, 29331824, 60424268, 29331825, 66714117, 29331826, 29331827, 35696052, 29331828, 33856870, 66712502, 52844045, 285007, 265009, 60433358, 264758, 55812038, 18108348, 52848317, 33109954, 33857084, 285017, 284604, 285018, 285019, 284882, 284369, 284288, 284688, 52844228, 21906768, 21906767, 21906768, 55811957, 285020, 285021, 285022, 52644150, 33857023, 52845129, 18108374, 18108378, 35898423, 58182323, 18108387, 87168518, 60432113, 22279000, 22279002, 284583, 284585
2875	88083726 (5749, 5750)	Novel Protein sim. GBank gij2781388 (AC004010) - similar to Leucine-rich transmembrane proteins; 44% similarity to U42767 (PID:8736918) [Homo sapiens]	Contains protein domain (PF00560) - glycoprotein Leucine Rich Repeat		22278996, 22278997, 22278999, 29331826, 29331828, 28146489, 66712502, 265008, 265017, 18108351, 284683, 284889, 21906767, 18108378, 18108377, 55811576, 80170394, 22279000, 284487
2876	88080854 (5751, 5752)	Novel Protein sim. GBank gij2979530 (AC004449) - R33683_2 [Homo sapiens]	Contains protein domain (PF00167) - Igt Fibroblast growth factor		52848365, 65274572, 58182575, 22278997, 22278998, 22278999, 60432048, 52645080, 60424269, 60432289, 29331827, 35696052, 29331828, 66712502, 52644045, 56182435, 60433358, 33657402, 33857084, 285019, 55811150, 284448, 284368, 21908766, 21908768, 21908769, 285020, 33857023, 33657109, 33657182, 27486262, 264628, 60431528, 55811576, 52844332, 56182323, 284558, 8337304, 18108385, 56526486, 60432113, 22279000, 22279002, 264583
2877	84747028 (5753, 5754)	Novel Protein sim. GBank gij4704208 (emb CAB41846.1) - (AL035419) dJ1100H1.1 (putative novel protein) [Homo sapiens]			264488, 284258, 28331822, 29331826, 264805, 284508, 284807, 284808, 284510, 285006, 284511, 284512, 33657402, 264758, 21906754, 18108351, 284881, 284882, 284288, 284684, 284685, 284768, 284689, 21908769, 284890, 33857023, 284693, 18108364, 33857109, 284829, 18108374, 284830, 284632, 284556, 284637, 284638, 284558, 18108385, 18108387, 284583, 284584, 284585, 284586, 284488, 284587
2878	88085308 (5755, 5756)	Novel Protein sim. GBank gij387675 (emb CAB03067) - (Z81077) predicted using GeneFINDER. Similarly to Yeast protein 8248 (TR:G587531) [Caenorhabditis elegans]	Contains protein domain (PF01207) - UNCLASSIFIED Uncharacterized protein family UPF0034		

2878	87880122 (5757, 5768)	Novel Protein sim. GBank gll4895145[gb AAD2752.1] - (AF127374) unknown [Streptomyces lavendulae]	Contains protein domain (PF00316) - Uracil-DNA glycosylase	UNCLASSIFIED	18106359, 264259, 264905, 18108370, 264629, 264908, 264909, 16106374, 18108377, 265008, 264910, 264637, 60170394, 264559, 265017, 264564, 264565, 264567, 264684, 264369
2880	94851439 (5758, 5760)	Novel Protein sim. GBank gll4880703[gb AAD2774.1]AF13298 - (AF132988) CGI-32 protein [Homo sapiens]			264488, 52846365, 52846842, 22278994, 35698266, 22278998, 22278999, 264259, 29331622, 29331624, 29331625, 29331626, 29331627, 35698052, 29331828, 264107, 264508, 264509, 264905, 264906, 264907, 264908, 264909, 52844045, 284510, 265006, 264511, 265007, 264512, 265009, 264910, 264594, 21908754, 52846317, 52844296, 87168558, 264600, 264604, 264605, 264760, 264764, 264288, 264768, 264769, 264687, 264769, 21908758, 21906769, 35695917, 265021, 264680, 264682, 33657023, 52845129, 33657108, 33657182, 27486262, 33657349, 264629, 16108374, 35695855, 264634, 264635, 264638, 264637, 264636, 284557, 52844332, 264558, 264559, 63373044, 284404, 22278900, 284563, 264483, 264567, 264466
2881	87850538 (5781, 5782)	Novel Protein sim. GBank gll7233571 (U23452) - No definition line found [Caenorhabditis elegans]		UNCLASSIFIED	22278998, 29331822, 52844045, 21906765, 264639, 60432113
2882	87714367 (5783, 5784)	Novel Protein sim. GBank gll1181112 (U41558) - No definition line found [Caenorhabditis elegans]			264488, 22278998, 22278999, 29331822, 29331826, 264908, 80170831, 60433356, 55612038, 264681, 264682, 264686, 264687, 264688, 21906768, 21906769, 264693, 263967, 18106374, 55611578, 56182323, 22279002, 264566
2883	95362675 (5785, 5786)	Novel Protein sim. GBank gll4688006[gb AAD31087.1]AF10693 - (AF106934) vitamin D receptor-interacting protein [Homo sapiens]	Contains protein domain (PF00400) - WD domain, G-beta repeat	UNCLASSIFIED	264112, 264662
2884	87784643 (5787, 5788)				265018, 264634
2885	83006308 (5789, 5790)	Novel Protein sim. GBank gll2224697[gb BA206321 - (ABD02378) KIAA0378 [Homo sapiens]		UNCLASSIFIED	284688, 264693
2886	91237623 (5771, 5772)	Novel Protein sim. GBank gll1255889 (U53344) - T07H8.5 gene product [Caenorhabditis elegans]	Contains protein domain (PF00084) - Sushi domain (SCR repeat)	complementrecept	60432049, 264259, 29331828, 264908, 264511, 264585, 60433438, 264596, 265017, 284605, 263969, 263972, 264555, 83373044, 87168518, 264568
2887	91227860 (5773, 5774)	Novel Protein sim. GBank gll366323[gb BA34521.1] - (AB016344) KIAA0801 protein [Homo sapiens]		UNCLASSIFIED	83373044, 264482
2888	95105618 (5775, 5778)	Novel Protein sim. GBank gll4508015[ref NP_003447.1]pZNF2 - zinc finger protein 205 Zinc finger, C2H2 type	Contains protein domain (PF00096) - Transcription factor		264488, 264259, 29331828, 264508, 264908, 264593, 264758, 264768, 264769, 16108374, 63373044, 264488

2889	87606562 (5777, 5778)			UNCLASSIFIED	56994075, 29331824, 265008, 264760, 18108354, 264288
2890	78703853 (5778, 5780)	Novel Protein sim. GBank g1854065[emb]CAAS83371 - (X83413) U88 [Human herpesvirus 8]		UNCLASSIFIED	264591, 264788
2891	88094428 (5781, 5782)	Novel Protein sim. GBank g113677750[emb]CAB015081 - (Z78064) predicted using GeneFinder, similar to collagen; cDNA EST EMBL:D65885 comes from this gene; cDNA EST EMBL:D89431 comes from this gene; cDNA EST EMBL:D66028 comes from this gene; cDNA EST EMBL:D69658 comes from this gene...		UNCLASSIFIED	264591, 264595, 264389, 264685, 264893, 264828, 264563, 264568
2892	95418745 (5783, 5784)	Novel Protein sim. GBank g14929759[emb]AA034140.3/AF15190 - (AF151903) CGH-145 protein [Homo sapiens]		UNCLASSIFIED	65274572, 35698286, 29331828, 264110, 265008, 60433438, 285018, 285019, 18108351, 264288, 21906785, 21908766, 21908769, 55811957, 264690, 65274620, 263987, 35695855
2893	87798014 (5785, 5786)			UNCLASSIFIED	264259, 285010, 264682, 18108370, 264555, 264558
2894	87755885 (5787, 5788)	Novel Protein sim. GBank g115669015[emb]AA048135.11 - (AF080171) zinc finger protein ZNF232 [Homo sapiens]	Contains protein domain (PF00096) - Zinc finger, C2H2 type	transcript factor	264259, 265008, 80433438, 52844298, 265011, 264389, 35695917, 18108381, 18108382, 18108388
2895	88938778 (5789, 5790)	Novel Protein sim. GBank g113824708[emb]CAAB6461 - (Z35597) Weak similarity with sea squirt nidogen precursor protein (blast score 71); cDNA EST EMBL:702089 comes from this gene; cDNA EST EMBL:D76135 comes from this gene; cDNA EST EMBL:D73147 comes from this gene; cDNA EST EMBL...	Contains protein domain (PF01437) - Plexin repeat		29331824, 265007, 264762, 264636, 264563
2896	87752122 (5791, 5792)	Novel Protein sim. GBank g1488554[emb]NP_005456.11pPRBG - protein kinase B gamma	Contains protein domain (PF00089) - Eukaryotic protein kinase domain	kinase	18108392, 18108394, 18108398, 22278988, 264259, 29331822, 29331824, 29148499, 264908, 264908, 265007, 265009, 265018, 265019, 264389, 284685, 264689, 21906766, 265021, 264893, 33857182, 284839, 18108384, 18108388, 264567
2897	95413057 (5793, 5794)	Novel Protein sim. GBank g14502877[emb]NP_001296.11pCLDN - Clostridium perfringens enterotoxin receptor 1		UNCLASSIFIED	60424179, 58182575, 22278998, 35896286, 22278997, 22278999, 80432049, 264259, 29331824, 29331825, 60424289, 60432289, 29331826, 29331828, 35696052, 264908, 58182435, 285009, 284910, 80170831, 60431735, 80433356, 80433438, 85274444, 55811388, 285018, 18108351, 284448, 284288, 264887, 21906785, 21908788, 21906787, 21906789, 55811957, 35895917, 264534, 33657023, 33857109, 35695763, 264828, 264629, 80431528, 18108374, 55810764, 55811578, 35698423, 35695855, 284555, 56182323, 18108385, 264404, 22279000, 22279002, 264566
2898	87750340 (5795, 5796)			UNCLASSIFIED	22278995, 264604, 18108385, 264566

2889	80357670 (5787, 5788)			UNCLASSIFIED	264764, 21906764, 264692
2900	94233538 (5789, 5800)	Novel Protein sim. GBank gll4581470[emb]CAB40137.1 - (Y16483) SLC7A8 protein [Homo sapiens]		glycoprotein	85274572, 58162575, 35896286, 60432049, 264259, 29331624, 66714117, 29331626, 35698052, 29331628, 68712502, 58162435, 285006, 265007, 265008, 285009, 60433356, 264736, 265018, 264764, 264785, 264286, 284788, 21908764, 21908768, 21908769, 265020, 264692, 264693, 32833986, 264631, 83373044, 60432113
2901	87444731 (5801, 5802)	Novel Protein sim. GBank gll4759272[ref]NP_004814.1[ptt]C4 - tetrairicopeptide repeat domain 4		phosphatase	22278995, 22278997, 22278998, 60432049, 29331822, 29331824, 29331825, 29331827, 35898052, 33856970, 284910, 265009, 21908754, 33657064, 67168474, 285010, 265018, 21906764, 21906785, 21908766, 21908767, 21908769, 33657023, 284683, 33657109, 33657349, 35896423, 35895655, 263981, 58182323, 22279002
2902	65745271 (5803, 5804)	Novel Protein sim. GBank gll2414815[emb]CAB16364 - (Z89259) hypothetical protein [Schizosaccharomyces pombe]			284683, 284691
2903	67808733 (5805, 5806)	Novel Protein sim. GBank gll1079318[ptt]S52241 - XLC12 protein - African clawed frog			284887, 22278994, 264259, 29331826, 29331828, 264905, 52644045, 58162435, 284511, 265017, 265018, 18108351, 284446, 284683, 264789, 264689, 35695917, 52644150, 67188518, 60432113, 22279002
2904	86458072 (5807, 5808)	Novel Protein sim. GBank gll5839823[gb]AA045885.1[AF14367 - (AF143676) multisplicing nuclear envelope membrane protein murim [Homo sapiens]		UNCLASSIFIED	52648365, 22278999, 264259, 35896052, 265011, 265017, 264883, 264789, 35695917, 265020, 263967, 18108374, 35895855, 284637, 264952, 18108365, 18108367
2905	84449826 (5809, 5810)	Novel Protein sim. GBank gll228837[sp]P39194[ALU7_HUMAN - III ALU SUBFAMILY SQ WARNING ENTRY III]		oncogene	265009, 284681, 284882
2906	95341051 (5811, 5812)	Novel Protein sim. GBank gll4689256[gb]AA027831.1[AF12185 - (AF121858) sorting nexin 8 [Homo sapiens]	Contains protein domain (PF00787) - UNCLASSIFIED Px domain		22278996, 35896286, 22278998, 264259, 60432289, 29331828, 29331830, 66712502, 285009, 60170631, 33109854, 264448, 264683, 264288, 264689, 21908766, 21908767, 21908768, 55811957, 35695917, 265022, 52644150, 264691, 33657023, 264692, 284693, 35895855, 60432113, 284598

2907	91211383 (5813, 5614)	Novel Protein sim. GBank g11707079 (U80451) - contains strong similarity to a DNA-like domain (P3:PS00636) [Caenorhabditis elegans]	Contains protein domain (PF00226) - eph DnaJ domain		52644507, 56182575, 56181686, 22278995, 56994075, 35696286, 80432049, 58182181, 35696052, 80431735, 284595, 55812038, 21906754, 55811386, 265018, 264682, 264369, 56181562, 21906768, 55811957, 35865917, 265020, 265021, 33657023, 33857109, 60431528, 55811578, 35898423, 35695855, 284638, 22279000
2906	80414246 (5815, 5816)	Novel Protein sim. GBank g12673917 (AC002561) - putative ATP-dependent RNA helicase [Arabidopsis thaliana]		helicase	265009, 33109954, 18108351, 264786, 265021, 264691, 264892, 18108374, 284558, 264638, 284557, 284558
2909	87420225 (5817, 5818)			eph	264259, 87188474, 265018, 18108365, 264628
2910	86601075 (5818, 5820)	Novel Protein sim. GBank g14533335 [Arabidopsis thaliana] (AL035539) putative protein [Arabidopsis thaliana]			22278995, 264509, 264512, 265007, 33857402, 285017, 284369, 285022, 18108365, 264628
2911	94216615 (5821, 5822)	Novel Protein sim. GBank g14468167 [emb] (CAB38415.1) - (AL031588) d1163J1.3 (novel protein similar to mouse B89) [Homo sapiens]		glucosylase	52646365, 18108397, 22278995, 22278997, 22278998, 22278999, 29331824, 29331825, 52644045, 265006, 265018, 284448, 21906765, 21906766, 21906767, 21906768, 21906769, 55811957, 265021, 18108370, 18108372, 18108374, 22279000
2912	8731803 (5823, 5824)	Novel Protein sim. GBank g14929837 [gb] (A034079.1) (AF151842) CGI-84 protein [Homo sapiens]	Contains protein domain (PF00904) - Involucrin repeat		52645158, 284092, 60432049, 264259, 52645080, 29331824, 29331825, 66712502, 33109954, 284760, 284883, 264288, 284886, 265021, 264693, 18108388, 263978, 264404
2913	87713823 (5825, 5828)	Novel Protein sim. GBank g1854063 [emb] (CAA58337) - (X63413) U88 [Human herpesvirus 8]	Contains protein domain (PF00099) - Zinc finger, C2H2 type	UNCLASSIFIED	52644507, 52645156, 52646842, 56182575, 35696286, 22278997, 264259, 52845080, 29331827, 35696052, 29331828, 264828, 52844045, 56182435, 55812038, 52848317, 21908754, 52644286, 87168474, 265017, 265018, 265019, 18108351, 264862, 264686, 264689, 21906765, 21906766, 21906767, 21908768, 35695917, 285020, 52844150, 27488281, 27488282, 27488283, 35695763, 55811578, 35695855, 52644332, 22279000, 22279002, 284583
2914	8787300 (5827, 5828)				264557

2915	88081972 (5828, 5630)	Novel Protein sim. GBank g15174465trfnp_006030.1ipfAA - endocytic receptor (macrophage mannose receptor family)	Contains protein domain (PF00059) - eph	264569, 264488, 264667, 264768, 21806766, 52646842, 21806767, 21906768, 56182375, 29146829, 35695917, 22278996, 22278997, 22278998, 265021, 22278999, 52644150, 264691, 264259, 60432049, 264892, 52645129, 33657109, 33657182, 29331827, 27486281, 35696052, 29331828, 27466262, 27486264, 27486265, 33657349, 29146498, 29146499, 264608, 264807, 18106370, 264508, 18106372, 52644045, 18106374, 56182435, 35695855, 264112, 264510, 265008, 60432229, 264593, 60433356, 56182323, 18106362, 55812038, 18108385, 33108854, 21906754, 33657064, 87168516, 265010, 265011, 60432113, 265017, 265018, 22278000, 265019, 55811150, 264881, 18106351, 264763, 264448, 264683, 264566, 18106354, 264389, 264288, 264788
2916	85337780 (5631, 5632)	Novel Protein sim. GBank g15104651db BAA0165.1 - (AP000061) 305aa long hypothetical GTP-4-dehydrohannose reductase [Aeropyrum pernix]	dehydrogenase	52645156, 85274572, 22278994, 22278995, 35686286, 22278996, 22278997, 22278998, 22278999, 264259, 29331822, 29331824, 29331825, 86714117, 60432288, 29331828, 29331827, 29331828, 33658970, 264509, 264806, 29331830, 52644045, 264909, 56182435, 80170831, 264592, 264593, 33657402, 60433356, 52646317, 21906754, 33108954, 33657084, 52644286, 65858542, 265011, 265017, 265018, 265019, 18108351, 264448, 264288, 52644229, 21906765, 21908766, 21906767, 21908768, 21906769, 55811857, 35695917, 265021, 265022, 52644150, 33657023, 33657109, 33657182, 27488281, 27488282, 27466264, 35695783, 16106376, 55811576, 35698423, 65274791, 35695855, 52644332, 264557, 264638, 56182323, 18106367, 67168516, 22278002, 264482
2917	87454546 (5833, 5834)	Novel Protein sim. GBank g1516805 emb CAA19260.1 - (AL023704) putative translocation elongation factor-Tu family [Schizosaccharomyces pombe]	UNCLASSIFIED	60433438, 264602, 264682, 87168518, 60432113
2918	85690528 (5635, 5836)	Novel Protein sim. GBank g1519216 pf S38038 - hypothetical protein YKL201c - yeast (Saccharomyces cerevisiae)		264838
2919	87641487 (5837, 5838)	Novel Protein sim. GBank g15264855 (AF030001) - unknown [Mus musculus]		86714117, 66712502, 263981

2820	87789523 (5838, 5840)					35696286, 22278997, 264259, 52645060, 29331824, 29331828, 29331827, 264628, 264908, 56182435, 264511, 264758, 33109954, 21908754, 52844298, 265010, 265011, 264601, 265017, 265018, 264681, 264687, 21908787, 285021, 52644150, 264680, 264681, 264682, 264693, 33657108, 33657182, 27466282, 27466284, 27466285, 35696423, 35695855, 284832, 264636, 264637, 264638, 56182323, 60170394, 18108385, 67188518, 60432113
2821	91639862 (5641, 5642)	Novel Protein sim. GBank gij45800138bIAAD24202.1 (U83194) TRAF4-associated factor 2 [Homo sapiens]	Contains protein domain (PF00787) - PX domain			35696286, 22278997, 264091, 264082, 264094, 264259, 29331822, 29331824, 29331826, 29331827, 35696052, 29146498, 264104, 264105, 264107, 264509, 264110, 264112, 264512, 60433358, 21906754, 87168474, 265017, 18108351, 264286, 21906785, 21908768, 21908787, 21908789, 35695917, 265021, 263974, 18108374, 283978, 283977, 18108376, 284555, 283981, 96526486, 67168518, 22279000, 22279002
2822	87749762 (5843, 5844)	Novel Protein sim. GBank gij4589514dbjBAA78779.11 - (AB023152) KIAA0935 protein [Homo sapiens]	Contains protein domain (PF01074) - Glycosyl hydrolases family 38		Kinase	264506, 264508, 264511, 285006, 285008, 284593, 33657402, 60174639, 18108351, 264763, 21906765, 29148627, 35695917, 264892, 264628, 263976, 56811578, 35695855, 284555, 284558, 56182323, 60170394, 22279000, 284486
2823	95337789 (5845, 5846)	Novel Protein sim. GBank gij4835268jembjCAB42898.21 - (Z83644) DJTIE18.4 (analog to mouse p116Rip protein) [Homo sapiens]	Contains protein domain (PF00169) - PH domain		struct	264488, 18108397, 22278995, 22278996, 22278997, 22278998, 22278999, 29331825, 29331828, 29331827, 29331830, 264511, 265009, 33657402, 285011, 285017, 265018, 284683, 18108354, 21906765, 21906767, 21906768, 21908768, 52844150, 264691, 264692, 33657109, 263974, 18108378, 264631, 264636, 18108385, 18108387, 22279000, 284583, 284568
2824	87791987 (5847, 5848)	Novel Protein sim. GBank gij2133085pjiS72254 - ribosomal protein L38, mitochondrial - yeast (Saccharomyces cerevisiae)	Contains protein domain (PF00444) - Ribosomal protein L38		ribosomal protein	265017, 264628, 20281152, 264558
2825	95090120 (5849, 5850)	Novel Protein sim. GBank gij23869898jembjCAB117181 - (Z68860) actin associated protein [Schizosaccharomyces pombe]			UNCLASSIFIED	56182575, 35686288, 264259, 60432288, 29331827, 264508, 52844045, 264910, 264591, 60432228, 55812038, 21906754, 264681, 264448, 264683, 264288, 264685, 52644228, 264689, 21906765, 21906788, 21906789, 21908789, 285021, 265022, 60170615, 264692, 33657023, 264693, 33657109, 35686423, 65274781, 56182323

2926	95343003 (5851, 5852)	Novel Protein sim. GBank glj283032lpn[IS22456 - hydroxyproline-rich glycoprotein - perinatal testis]				29331826, 265011, 264768, 264689
2927	80408018 (5853, 5854)	Novel Protein sim. GBank glj3378056 (AF017777) - helicase				284764, 264288, 264630, 264637
2928	20452179 (5855, 5856)	Novel Protein sim. GBank glj3413320[emh(CAA06915) - (A)069215] CMP-N-acetylneuraminic acid synthetase [Mus musculus]		UNCLASSIFIED		264559 264569, 264469, 22270994, 35696286, 22276996, 22278999, 264259, 29331825, 264259, 52645080, 29331622, 29331624, 264259, 52645080, 29331622, 29331624, 66714117, 29331825, 29331826, 29331827, 35696052, 33856970, 264108, 29331630, 52644045, 265009, 33109954, 52644286, 87168559, 264760, 264762, 264448, 264784, 264288, 264766, 264768, 21908765, 21906786, 21906768, 21906769, 35695917, 264691, 33657023, 264883, 33657109, 18108374, 263976, 35698423, 35695855, 263981, 22279000, 22279002, 264567, 264438
2930	95302755 (5859, 5860)			UNCLASSIFIED		56182575, 56181666, 35696286, 22276996, 22278998, 22278999, 264259, 29331825, 60432288, 29331628, 264905, 52644045, 56182435, 285009, 60170631, 264592, 60432229, 60433356, 87168474, 265010, 265011, 265017, 265016, 265019, 264762, 264448, 264883, 264288, 264766, 21906785, 21906769, 35695917, 60170815, 33657023, 33657109, 264628, 18108370, 18108372, 35696423, 35695855, 264556, 56182323, 60432113, 264567
2931	94312693 (5861, 5862)	Novel Protein sim. GBank glj3786433 (AF098605) - similar to Arabidopsis thaliana male sterility protein 2 (SW-C08891) [Caenorhabditis elegans]	Contains protein domain (PFO0471) - UNCLASSIFIED Ribosomal protein L33			52645156, 22278997, 22278998, 29331822, 52645080, 29331824, 60432288, 33656970, 60433356, 60433438, 33109954, 21906765, 21906766, 21906767, 21906768, 265020, 52644150, 33657023, 33657109, 33657182, 27468265, 35698423, 35695855, 264555, 67168518, 60432113, 264568
2932	79632623 (5863, 5864)					264906, 264907
2933	81720776 (5865, 5866)	Novel Protein sim. GBank glj3378056 (AF017777) - helicase [Drosophila melanogaster]		helicase		264486, 16108382, 56182575, 22276999, 264091, 264259, 29331825, 60432288, 29331827, 264506, 52644045, 56182435, 265007, 265009, 264592, 60433356, 60433438, 21906754, 265017, 264682, 264288, 52644229, 21906765, 21906766, 21906768, 21906769, 265022, 52644150, 33657023, 33657109, 27468265, 264835, 264636, 60170394, 56182323, 16108385, 60432113, 264565, 264568, 264587

2934	86576025 (5867, 5868)				22278997, 22278999, 29331824, 33657402, 264691, 27486282, 264628, 87166516, 22279000
2935	86410578 (5868, 5870)			UNCLASSIFIED	56182575, 22278995, 60433356, 33657402, 264758, 33109954, 21906754, 265016, 265019, 264448, 264769, 21906764, 21906765, 265021, 264692, 33657023, 33657109, 33657349, 55810784, 22279000, 22278997, 29331827, 29331828, 265009, 265017, 264605, 265020, 55811578, 18108387, 60432113, 264563
2938	87805853 (5871, 5872)	Novel Protein sim. GBank gij4133682 (AC005065) - determined by GENSCAN prediction and spliced EST: match to EST R84329 (MID:942735) [Homo sapiens]	Contains protein domain (PF00856) - SET domain	nuclease	56984075, 22278999, 264259, 60432049, 29331822, 56182161, 29331827, 29331828, 264906, 264906, 264909, 56182435, 265006, 264512, 264910, 60170831, 60433356, 265011, 265018, 18108351, 264448, 264288, 264768, 52644229, 21906765, 29148784, 85274791, 264558, 56162323, 60170394, 264558, 60432113, 264585, 264488, 264587, 264488, 56182575, 22278998, 33686286, 22278997, 22278998, 22278999, 264480, 264259, 29331822, 29331824, 66714117, 29331827, 35698052, 264107, 264505, 66712502, 52644045, 56182435, 264511, 265008, 265009, 60432229, 33657402, 60433438, 55812038, 21906754, 65858542, 265010, 265011, 87166558, 265017, 265016, 265019, 264681, 264286, 264689, 21906765, 21906767, 21906768, 55811957, 35695917, 265020, 60170815, 264690, 264691, 264692, 33657023, 264693, 65274620, 33857109, 18108370, 18108374, 263978, 35698423, 35695855, 264555, 264558, 18108381, 56182323, 60170394, 83373044, 18108365, 56526486, 60432113, 22279002, 264907, 265016, 264681, 264685, 264686
2938	95419773 (5875, 5876)	Novel Protein sim. GBank gij3318980 (embICA7672Q) - (Y17267) ubiquitin-conjugating enzyme [Mus musculus]	Contains protein domain (PF00170) - Ubiquitin-conjugating enzyme	ubiquitin	264488, 56182575, 22278998, 33686286, 22278997, 22278998, 22278999, 264480, 264259, 29331822, 29331824, 66714117, 29331827, 35698052, 264107, 264505, 66712502, 52644045, 56182435, 264511, 265008, 265009, 60432229, 33657402, 60433438, 55812038, 21906754, 65858542, 265010, 265011, 87166558, 265017, 265016, 265019, 264681, 264286, 264689, 21906765, 21906767, 21906768, 55811957, 35695917, 265020, 60170815, 264690, 264691, 264692, 33657023, 264693, 65274620, 33857109, 18108370, 18108374, 263978, 35698423, 35695855, 264555, 264558, 18108381, 56182323, 60170394, 83373044, 18108365, 56526486, 60432113, 22279002, 264907, 265016, 264681, 264685, 264686
2939	87786822 (5877, 5878)	Novel Protein sim. GBank gij3978900 (embICA93909) - (Z75547) similar to WD domain, G-beta repeat; cDNA EST YK371b7.5 comes from this gene; cDNA EST YK312n1.5 comes from this gene; cDNA EST YK465d5.5 comes from this gene; cDNA EST YK472c4.5 comes from this gene; cDNA EST YK292i8. ...	Contains protein domain (PF00400) - WD domain, G-beta repeat	ATPase-associated	

2840	95011103 (5879, 5880)			UNCLASSIFIED	22278996, 29331822, 29331824, 66714117, 29331826, 29331828, 264905, 264908, 66712502, 29331830, 265011, 265017, 264764, 264369, 21906766, 21906767, 33657023, 33657109, 32833966, 18108374, 18108377, 264634, 63373044, 18108385, 18108387, 264566
2841	21423370 (5881, 5882)	Novel Protein sim. GBank gij3413672p[BA32300] - (AB007924) KIAA0455 protein (Homo sapiens)		UNCLASSIFIED	284557
2842	87430203 (5883, 5884)	Novel Protein sim. GBank gij1172845p[46629R825_RABIT - RAS-RELATED PROTEIN RAB-25]		glycoprotein	264910, 265010, 264768
2843	85314504 (5885, 5886)	Novel Protein sim. GBank gij4929653ig[PA034067_1(A-F15185 - (A-F151850) CG1-92 protein (Homo sapiens)]		collagen	60432049, 284259, 60432288, 29331827, 29140496, 265008, 264593, 80433356, 80433438, 265010, 265011, 265017, 265018, 264683, 264766, 18108381, 65274727, 60432113, 264567
2844	85081063 (5887, 5888)	Novel Protein sim. GBank gij678282p[CB41190_1] - (AL049680) 1-acetyl-3-phosphate acyltransferase-like protein (Arabidopsis thaliana)	Contains protein domain (PF00415) - Regulator of chromosome condensation (RCC1)	ATPase-associated	56984075, 22278998, 60432049, 284259, 29331822, 29331824, 60424268, 60432289, 29331828, 29331828, 264905, 284907, 52644045, 264908, 264511, 265008, 265008, 264594, 21906754, 87168558, 284603, 265017, 265018, 18108351, 264882, 284766, 264687, 264689, 21906765, 21906766, 21906767, 21906768, 21906769, 265021, 60170615, 52644150, 264690, 264691, 33657023, 264692, 264693, 33657109, 33657182, 33657349, 18108370, 18108374, 18108377, 55811578, 35698423, 35698555, 264835, 264555, 264556, 56182323, 60170394, 264558, 264559, 83373044, 58526486, 87188518, 60432113, 22279002, 264482, 264563, 264484, 264567
2845	94233560 (5889, 5890)	Novel Protein sim. GBank gij726831p[P39188]ALU1_HUMAN - III ALU SUBFAMILY J WARNING ENTRY III	Contains protein domain (PF00096) - Zinc finger, C2H2 type	UNCLASSIFIED	60424179, 22278995, 22278996, 22278998, 22278999, 264259, 58182181, 29331824, 22278999, 60432288, 35696052, 264908, 60424268, 60433356, 55612038, 264759, 265008, 60433356, 55612038, 264759, 55811386, 285018, 284681, 18108351, 284446, 264683, 264389, 264288, 264687, 56181562, 21906767, 21906768, 21906769, 35695917, 265020, 265021, 264693, 60431528, 55610764, 35698423, 35695855, 264630, 60170394, 63373044, 22279000, 264588, 264567

2848	84317315 (5891, 5892)	Novel Protein sim. GBank gij5441952gpiAAC43195.1(AFO7286 - (AF072864) peroxisomal membrane protein PMP 24 [Homo sapiens]	UNCLASSIFIED	264488, 264259, 264508, 264509, 264908, 264907, 264909, 264510, 264511, 265007, 264512, 264910, 264591, 264593, 18108351, 264764, 264288, 264684, 264789, 265021, 264892, 33857109, 264628, 264829, 18108374, 264631, 284634, 264636, 264637, 18108380, 264638, 264639, 63373044, 264565, 264566, 264486, 264587
2847	87362952 (5893, 5894)	Novel Protein sim. GBank gij3540281gpiAAC34383.1 - (AF058118) Ali-1 related protein [Fugu rubripes]	UNCLASSIFIED	22278995, 22278996, 22278997, 22278998, 29148498, 264508, 29331830, 265007, 265008, 265009, 80432228, 21908754, 265010, 265017, 265019, 264766, 264685, 21908785, 21908786, 21908787, 21908788, 21908789, 265020, 264628, 18108370, 264629, 284830, 18108387, 60432113
2846	87826527 (5895, 5896)	Novel Protein sim. GBank gij5566514gpiAB65654.2 - (AF001533) mitogen-induced [Mus musculus]		5264642, 22278995, 264259, 29331824, 29331825, 29331827, 29331830, 284909, 265007, 265009, 265019, 264763, 264684, 264288, 264685, 264686, 21908787, 264691, 264692, 264693, 18108374, 55811578, 18108385, 22279002, 264563, 264567
2849	88175545 (5897, 5898)	Novel Protein sim. GBank gij2132923gpiS67133 - probable membrane protein YOR240w - yeast [Saccharomyces cerevisiae]	UNCLASSIFIED	22278998, 22278997, 60432289, 29331826, 29331827, 29331828, 35696052, 29146499, 264104, 264107, 264905, 88712502, 264808, 60433356, 60433438, 87168559, 264764, 52644229, 58181562, 21908787, 21908788, 21908789, 265022, 60170615, 33857023, 35696423, 263981, 264558, 60432113, 22279002
2850	95088870 (5899, 5900)	Novel Protein sim. GBank gij4661025gpiP34629YCO8_CAEEL - PUTATIVE AMINOPEPTIDASE ZK353.8 IN CHROMOSOME III	Contains protein domain (PF00883) - peptidase Cytosol aminopeptidase family	264488, 35696286, 264259, 35696052, 264907, 265007, 264910, 265017, 265018, 264288, 264788, 35695917, 265020, 18108362, 18108370, 18108378, 35696423, 65274791, 35695855, 264558, 58528488, 264486
2851	87392357 (5901, 5902)	Novel Protein sim. GBank gij4688025gpiCAB41450.1 - (AJ238248) centaurin beta2 [Homo sapiens]		264893

2952	95329952 (3903, 5904)	Novel Protein sim. GBank g1558663jemhCAE51405.11 - (AL096881) hypothetical protein [Homo sapiens]	Contains protein domain (PF00630) - CRL/TRIO domain.	transcript factor	264887, 52845156, 21808766, 21808769, 22278896, 265020, 264680, 60432049, 264259, 264693, 29331822, 16108385, 29331825, 80432289, 33657108, 18108368, 29331827, 35696052, 27486262, 264508, 264905, 20281149, 264906, 264907, 29331830, 264908, 264909, 35895855, 264511, 265008, 265009, 264910, 264635, 264836, 60432229, 264638, 60433358, 264639, 264758, 67166518, 265017, 22279000, 22278002, 264760, 264583, 264482, 18108351, 264448, 264288
2953	88093575 (5805, 5906)	Novel Protein sim. GBank g11195225p106585ERC_RABIT - PROBABLE PHOSPHOESTERASE AMINOTRANSFERASE (PSAT) (ENDOMETRIAL PROGESTERONE-INDUCED PROTEIN) (EPIP)	Contains protein domain (PF00288) - Amino transferases class-V	UNCLASSIFIED	18106388, 56984075, 22278896, 29331822, 29331824, 29331825, 29331827, 35696052, 264508, 264905, 264908, 264907, 264510, 264591, 264594, 33657402, 264585, 264586, 264758, 52846317, 21808754, 33657084, 52844298, 67166559, 264600, 264780, 264661, 18108351, 264764, 264389, 264288, 264887, 21906785, 21908768, 21906787, 21906768, 21906769, 35695917, 33657023, 16108384, 52845128, 33657109, 33657349, 16108374, 263978, 35695655, 264837, 264638, 67166518, 264482, 264583, 264585, 29331822
2954	86086288 (5907, 5908)	Novel Protein sim. GBank g14865281p106585 - growth differentiation factor 9	Contains protein domain (PF00018) - transforming growth factor beta like domain	igf	264259, 29331822, 29331824, 29331825, 29331826, 35696052, 264608, 52844045, 264512, 60432229, 285016, 265018, 55811150, 264769, 21908767, 21908768, 21908789, 265021, 60170815, 55610764, 264567
2955	87686428 (5908, 5910)	Novel Protein sim. GBank g13452473 (AF084205) - serine/threonine protein kinase TAO1 [Rattus norvegicus]	Contains protein domain (PF00069) - Eukaryotic protein kinase domain	kinase	22278996, 264259, 29331827, 264908, 21906786
2956	65786745 (5911, 5912)	Novel Protein sim. GBank g14869254p106585 - (AF121857) sorting nexin 7 [Homo sapiens]	Contains protein domain (PF00767) - PX domain		22278996, 264259, 29331824, 29331827, 29331829, 264595, 264758, 265010, 265011, 264448, 264763, 264683, 264288, 264685, 18108357, 29148829, 264890, 18108362, 264893, 16108370, 60431528, 18108374, 264634, 18108381, 56162323, 18108382, 18108385, 16106386, 56526488, 67168518, 264462, 264487
2957	80933301 (5913, 5914)	Novel Protein sim. GBank g14503023p106585 - 000089, 1PCP17 - carnitine palmitoyltransferase II precursor		cadherin	264595, 264596, 264681, 264389, 264629, 264631, 264587
2958	87440014 (5915, 5916)	Novel Protein sim. GBank g14240257p106585 - KIAA0884 protein [Homo sapiens]			

2989	85109420 (5917, 5918)	Novel Protein sim. GBank gll596646lemj(CAB05177.2) - [musculus]	Contains protein domain (PF00566) - oncogene	263994, 22278997, 264259, 60432049, 29331826, 29331828, 35696052, 29331830, 6671502, 58182435, 285006, 264512, 265008, 285009, 60433356, 60433438, 264596, 285017, 285018, 264683, 264288, 264768, 264769, 21906766, 21906767, 21906769, 265020, 80170615, 264892, 27486285, 18108374, 65274791, 35695855, 8337304, 56526486, 60432113
2990	87420081 (5918, 5920)		UNCLASSIFIED	35696286, 56182435, 87188474, 285010, 80170815, 35698423, 58182323, 18108383, 87188518, 284483
2981	85413416 (5921, 5922)	Novel Protein sim. GBank gll596646lemj(CAB05177.2) - (Z82288) predicted using GeneFINDER: similar to WD domain, G-beta repeats [Caenorhabditis elegans]	Contains protein domain (PF00400) - Transcription factor	22278997, 22278999, 264259, 29331822, 29331824, 29331828, 29331829, 284907, 264908, 52644045, 265006, 33857402, 21906754, 87166474, 285011, 87188559, 265017, 21908769, 265020, 60170815, 264692, 33857023, 35695763, 18108370, 16108374, 35698423, 264632, 264838, 18108385, 87188518, 22279002, 264584, 264587
2982	87912700 (5923, 5924)		UNCLASSIFIED	35696286, 22278997, 264092, 264094, 264259, 29331824, 66714117, 29331825, 80432288, 29331828, 29331827, 29331828, 35695052, 264508, 264905, 264509, 264907, 264908, 264909, 264510, 264512, 264593, 264594, 60433438, 264736, 52646317, 264602, 264803, 264605, 264760, 264762, 264764, 264288, 264766, 264686, 264788, 264769, 35695917, 265020, 264691, 264634, 264638, 264837, 264638, 264838, 18108385, 264583, 264585, 264586, 264587, 264488
2983	89313464 (5925, 5926)	Novel Protein sim. GBank gll42402231dbj(BAA74890.1) - (AB020674) KIAA0867 protein [Homo sapiens]	Contains protein domain (PF00010) - Helix-loop-helix DNA-binding domain	18108392, 56994075, 22278998, 22278999, 29331822, 29331825, 29331828, 29331827, 29331828, 265007, 265008, 264592, 264594, 21908754, 285018, 284780, 264687, 29148627, 29148784, 265020, 33857023, 284893, 85274820, 33857182, 27488261, 264628, 55810764, 35696423, 284555, 284838, 264837, 284557, 264538, 264563
2984	84324817 (5927, 5928)		UNCLASSIFIED	284259, 29331828, 33857402, 265017, 265018, 264692, 18108388, 35698423, 83373044, 18108388

2965	80364762 (5929, 5930)	Novel Protein sim. GBank gi 4685447 ref NP_005452.1 pKRM1 - Keisler (mouse) mal-related leucine zipper homolog		transcription factor	264259, 29331828, 264506, 264508, 264509, 264511, 265008, 264907, 264908, 264909, 264591, 264593, 264594, 33657402, 264910, 264591, 264593, 264594, 33657402, 265011, 264780, 264762, 264764, 264288, 264805, 264786, 264892, 33657109, 264628, 264828, 35685655, 264630, 264831, 264632, 264634, 264835, 264638, 264637, 264636, 264839, 264583, 264587, 18108391, 80432288, 264682, 264446
2966	91725248 (5831, 5832)	Novel Protein sim. GBank gi 5282751 am01CAB45690.1 - (A)Z43177) Xenopus RPA interacting protein alpha [Xenopus laevis]			
2967	94658303 (5933, 5934)	Novel Protein sim. GBank gi 624225 (U19181) - Rabn3 [Rattus norvegicus]		UNCLASSIFIED	264488, 264506, 264509, 264808, 264909, 264511, 264810, 264594, 264738, 85658542, 264782, 264764, 265021, 264556, 18108381, 264584, 264486
2968	95302776 (5835, 5836)	Novel Protein sim. GBank gi 4929715 gb AA034118.1 AF15188 - (AF151881) CGI-123 protein [Homo sapiens]	Contains protein domain (PF00097) - Zinc finger, C2H2 type (RING finger)		264687, 52645156, 21906765, 52646365, 21908767, 18108398, 35686423, 22278996, 35686266, 22278997, 285020, 22278999, 265021, 265022, 264093, 264836, 264690, 52644150, 264238, 33657023, 52645080, 264693, 29331822, 58182181, 29331824, 66714117, 29331825, 33109954, 52645129, 29331828, 21906754, 33657182, 29331827, 29331828, 35686052, 27488262, 87168518, 87168474, 285010, 87168559, 265018, 22278000, 265018, 22279002, 264563, 18108351, 264806, 264807, 264448, 66712502, 264568, 264369, 264288
2969	95310957 (5937, 5938)	Novel Protein sim. GBank gi 3024743 p LOC47347HSA_SULST - THERMOSOME, ALPHA SUBUNIT (CHAPERONIN ALPHA SUBUNIT)		epn	52646842, 22278898, 22278898, 22278899, 60432046, 264258, 29331824, 29331825, 29331828, 29331828, 264508, 264808, 52644045, 56182435, 285009, 80433438, 55812036, 21906754, 285011, 87168559, 265018, 265019, 264446, 284288, 264389, 52644228, 21906766, 21906768, 21906769, 29148784, 265020, 285021, 52644150, 284691, 33657109, 18108374, 58182323, 60170394, 87168518, 80432113, 22278000
2970	88088071 (5939, 5940)	Novel Protein sim. GBank gi 3165407 (AC004755) - (os37502_1 (Homo sapiens)	Contains protein domain (PF00046) - Homeobox domain	homeobox	

2971	94196930 (5941, 5942)	Novel Protein sim. GBank gij728837ispf39194IALUT_HUMAN - III ALU SUBFAMILY SQ WARNING ENTRY III		lm7	264488, 56182575, 35696266, 56994075, 29331824, 29331828, 29148489, 264508, 264905, 264907, 264112, 264910, 21908754, 67168559, 265018, 265019, 18106351, 264689, 21906765, 21906767, 21906768, 265020, 265021, 60170815, 18108364, 264828, 264829, 18108374, 264636, 264556, 264558, 83373044, 18108384, 18108385, 67168518, 264594, 264597 265017, 35695917, 265021, 33657109, 22279002, 264563
2972	88625943 (5943, 5944)	Novel Protein sim. GBank gij728838ispf39193ALUe_HUMAN - III ALU SUBFAMILY SP WARNING ENTRY III		kinase	29331822, 264692, 33657349, 55811576, 264563
2973	91215301 (5945, 5946)	Novel Protein sim. GBank gij2746789 (AF040842) - No definition line found [Caenorhabditis elegans]		UNCLASSIFIED	29331822, 264692, 33657349, 55811576, 264563
2974	91673002 (5947, 5948)	Novel Protein sim. GBank gij786117 (L41634) - nuclear protein [Ensis minor]		UNCLASSIFIED	18108392, 52644507, 56182575, 58181686, 22278995, 22278998, 35696266, 22278997, 22278996, 22278999, 264259, 52845080, 29331824, 29331825, 66714117, 60424289, 29331828, 29331827, 29331828, 35696052, 66712502, 264908, 52644045, 265007, 264910, 265008, 80433438, 33109954, 21906754, 55811386, 52644286, 87166474, 87168559, 265017, 265018, 265019, 18108351, 264448, 264369, 264288, 52644229, 18108359, 21906785, 21906787, 21906768, 35695917, 265020, 265021, 52644150, 264691, 264692, 33657023, 27466282, 27466284, 35695763, 18108370, 18108376, 55810764, 55811576, 35696423, 35695855, 264630, 264635, 264557, 52644332, 264556, 83373044, 18108387, 87168518, 80432113, 22279000, 264482, 264467
2975	95325213 (5949, 5950)	Novel Protein sim. GBank gij3880812embCAA195081 - (AL023839) similar to HECT-domain (ubiquitin-transferase); cDNA EST YK480410.5 comes from this gene	Contains protein domain (PF00632) - ubiquitin- HECT-domain (ubiquitin- transferase).	ubiquitin	29331824, 29331827, 29331628, 264910, 65658542, 265011, 265018, 264448, 264288, 264769, 21908787, 265020, 264691, 264559, 83373044
2976	87771202 (5951, 5952)	Novel Protein sim. GBank gij5678138bbJAD48674.1JAF16093 - (AF160934) BcDNA.LD14189 [Drosophila melanogaster]		transport	22278998, 264908, 265007, 265010, 265011, 265017, 265019, 18106351, 264685, 264689, 18108370, 264639, 18108385
2977	91725254 (5953, 5954)	Novel Protein sim. GBank gij5262751embCAB45690.11 - (AJ243177) Xenopus RPA interacting protein alpha [Xenopus laevis]		UNCLASSIFIED	264509, 264288

2978	67332059 (5955, 5956)	Novel Protein sim. GBank gll746549 (U23522) - No definition line found (Caenorhabditis elegans)	Contains protein domain (PF00480)	UNCLASSIFIED	22278995, 22278998, 22278997, 22278999, 264259, 60432289, 29331827, 29146499, 56182435, 285008, 265007, 265009, 60433356, 60433438, 21906754, 285010, 265011, 265017, 265018, 265019, 264288, 264685, 264688, 21906765, 21906766, 21906767, 21906768, 21906769, 265020, 265021, 265022, 35896423, 264639, 60432113, 22279000, 22279002
2979	91725258 (5957, 5958)	Novel Protein sim. GBank gll5262751 (emb)CABA5890.11 - (AJ243177) Xenopus RPA interacting protein alpha (Xenopus laevis)		complement	264486, 65274572, 56994075, 22278999, 264093, 29331822, 29331824, 264288, 55811957, 33657023, 33657109, 18108370, 55811578, 58182323, 60432113, 284482
2980	68296600 (5959, 5960)				265009, 21906767, 283981, 22279000
2981	67376330 (5961, 5962)	Novel Protein sim. GBank gll4928767 (gb)AAD34144.1 (AF151907) CGI-149 protein (Homo sapiens)		UNCLASSIFIED	264628, 264564
2982	95303675 (5963, 5964)				22278995, 56994075, 22278998, 22278997, 22278999, 22278999, 264092, 29331824, 29331827, 29331828, 264905, 284591, 264592, 264594, 264595, 264596, 33657084, 264448, 21906765, 21906766, 21906767, 21906768, 21906769, 265020, 265022, 18108365, 33657162, 33657349, 35688423, 63373044, 22279000, 22279002
2983	91725258 (5965, 5966)	Novel Protein sim. GBank gll5262751 (emb)CABA5690.11 - (AJ243177) Xenopus RPA interacting protein alpha (Xenopus laevis)			60424178, 52848842, 18108398, 22278997, 264093, 60432049, 264259, 29331822, 60432289, 33858870, 284805, 52844045, 265006, 60431735, 67168474, 265018, 265019, 18108351, 264448, 21906765, 21906766, 35895917, 33657023, 52845129, 18108370, 35896423, 63373044, 56528488, 60432113, 264404, 22279002
2984	94136467 (5967, 5968)	Novel Protein sim. GBank gll2393734 (AC002542) - similar to C. elegans F11A10.5; 80% similarity to Z68297 (FIDg1130618) (Homo sapiens)	ATPase-associated		
2985	67098072 (5969, 5970)	Novel Protein sim. GBank gll103180 (pili)S22128 - finger protein unkempt - fruit fly (Drosophila melanogaster)		UNCLASSIFIED	264910, 55612038, 56181562, 55611957, 284628, 55810764, 284632, 284635, 60432113
2986	68284861 (5971, 5972)				55611957, 264566
2987	86455834 (5973, 5974)			UNCLASSIFIED	284369

2988	95357153 (5975, 5976)	Novel Protein sim. GBank g146190261bAAAD27002.11 - (AF071207) HSPC021 [Homo sapiens]		UNCLASSIFIED	264488, 85274572, 22278995, 22278998, 22278997, 22278999, 264092, 264094, 284259, 80432049, 2831624, 2931628, 60432289, 35698052, 2931628, 284107, 284905, 264807, 264808, 66712502, 284828, 264808, 56182435, 285006, 285007, 265008, 60170831, 60432229, 264593, 60433358, 264757, 60433438, 21908754, 265010, 265011, 87188559, 285017, 265018, 264682, 264448, 264369, 264288, 264685, 52644229, 21906765, 21908767, 21908769, 35695917, 265021, 265022, 52644150, 264890, 33857023, 65274820, 263867, 33657109, 27486262, 18108370, 18108372, 18108374, 55810764, 65274791, 35695855, 284635, 264836, 264637, 263961, 264638, 58183232, 83373044, 60432113, 22279000, 264563, 264564, 264585, 264568, 264567, 22278998, 22278997, 264905, 264511, 80170831, 264593, 265019, 21908765, 21906767, 21908768, 18108374
2989	81225118 (5977, 5978)	Novel Protein sim. GBank g11138711spIP23984ALUF_HUMAN - III ALU CLASS F WARNING ENTRY III		kinase	264563, 264564, 264585, 264568, 264567, 22278998, 22278997, 264905, 264511, 80170831, 264593, 265019, 21908765, 21906767, 21908768, 18108374
2990	67330444 (5979, 5980)	Novel Protein sim. GBank g126296381spIP973481RHOD_MOUSE - RHO-RELATED GTP-BINDING PROTEIN RHOD	Contains protein domain (PF00071) - oncogene	UNCLASSIFIED	265007, 264512, 18108351, 264288, 264689, 265020, 264891, 33657023, 33657109
2991	84325361 (5981, 5982)			UNCLASSIFIED	264563
2992	85425164 (5983, 5984)			UNCLASSIFIED	264259, 285019, 264689, 18108385, 264488, 29331822, 285017, 264761, 21906769, 85274791, 263981, 284555
2993	94325363 (5985, 5986)			UNCLASSIFIED	22278994, 22278995, 56894075, 22278997, 22278999, 264259, 29331822, 29331824, 29331825, 29331826, 29331827, 29331828, 265008, 265009, 264910, 33109954, 87168474, 87168559, 265018, 285019, 264448, 264288, 21906768, 21906767, 21906769, 265021, 265022, 33857023, 264693, 35695855, 83373044, 18108385, 22279000, 264565, 264566, 264905, 264907, 285019, 18108351, 264683
2994	94138834 (5987, 5988)	Novel Protein sim. GBank g124865491spID506591YU02_MYCTU - HYPOTHETICAL 28.7 KD PROTEIN CY339.02		transport	65274572, 35698286, 264259, 29331824, 35696052, 29146499, 264508, 284907, 285007, 285008, 60433436, 18108348, 265017, 264681, 264683, 284288, 264766, 284769, 264689, 35695917, 80170615, 33657023, 264692, 264634, 264555, 18108381, 18108382, 18108389, 264464
2995	67591070 (5989, 5990)	Novel Protein sim. GBank g12734081 (AF000195) - similar to oxygen-binding proteins [Caenorhabditis elegans]		helicase	
2996	81013786 (5991, 5992)	Novel Protein sim. GBank g12628912 (AC002281) - Similar ATP-dependent RNA Helicase [Arabidopsis thaliana]	Contains protein domain (PF00270) - DEAD/DEAH box helicase	helicase	

2897	87627440 (5893, 5994)	Novel Protein sim. GBank g14589852jdbj BA76848-1 - (AB023221) KIAA1004 protein [Homo sapiens]	homeobox	264488, 56182575, 284259, 68714117, 29331826, 35898052, 264508, 264509, 284907, 264908, 285008, 67168474, 265019, 284448, 284682, 264685, 264766, 21906764, 21906766, 21906768, 21906789, 27488281, 18108374, 35898423, 284634, 264635, 264638, 264557, 18108385, 87188518, 52845385, 22278997, 284508, 264808, 18108351, 21906785, 21906787, 18108370, 18108374, 35898423, 264638, 264639
2898	88085381 (5895, 5898)	Novel Protein sim. GBank g13547588jemb CAA22521 - (AL034364) cDNA EST yk255b8.3 comes from this gene: cDNA EST yk255b8.5 comes from this gene: cDNA EST EMBL.M75923 comes from this gene [Caenorhabditis elegans]	UNCLASSIFIED	52845385, 22278997, 284508, 264808, 18108351, 21906785, 21906787, 18108370, 18108374, 35898423, 264638, 264639
2899	94847055 (5987, 5989)	Novel Protein sim. GBank g11154083p P8835 CC19_CAEEL - CUTICLE COLLAGEN 18	UNCLASSIFIED	56182575, 22278996, 29147620, 28331825, 29146498, 29146499, 284905, 68712502, 265008, 265009, 21906754, 85658542, 18108351, 29148627, 29148628, 60170615, 33657109, 27486282, 18108370, 18108374, 264536, 264557, 264538, 60170394, 18108385, 284563
3000	95089370 (5999, 6000)	Novel Protein sim. GBank g11163174 (U32575) - similar to yeast Sec8p, Swiss-Pro Accession Number P32844; similar to mammalian B84, Swiss-Pro Accession Number Q03188; Method: conceptual translation supplied by author [Rattus norvegicus]	UNCLASSIFIED	264887, 22278997, 22278999, 284259, 29331822, 29331824, 35898052, 28146498, 264508, 264905, 264908, 284907, 264908, 264909, 284510, 284511, 265006, 285007, 265008, 265009, 264810, 33657402, 284757, 264595, 264596, 264758, 21906754, 285011, 284800, 265017, 265018, 264605, 265019, 264760, 264761, 264762, 264681, 264682, 264764, 264288, 264685, 264768, 264686, 264768, 284768, 21906765, 21906768, 35895917, 265020, 284691, 284692, 33657023, 284693, 33657109, 33657182, 27466261, 284628, 284629, 18108374, 18108376, 35898423, 35898655, 284630, 284631, 284632, 284634, 284635, 284638, 264637, 284556, 284638, 284639, 80170394, 63373044, 20788451, 22279002, 264563, 284488, 284567
3001	88078454 (6001, 6002)	Novel Protein sim. GBank g12078470 (AC002073) - Putative gene. Genscan predictions confirmed by EST splicing: coded for by human cDNAs AA122029 (NID:81878049), D31582 (NID:8844442), AA158721 (NID:81733515), R59840 (NID:8830335) and F13082 (NID:8708111) [Homo sapiens]	caltropsh	18108394, 52846842, 56182575, 29331824, 29331825, 29331827, 264910, 33108954, 52844288, 285017, 285019, 264288, 285020, 285021, 52844150, 284692, 35895763, 55810764, 35898423, 56182323, 18108387, 284488, 284564
3002	87718167 (6003, 6004)	Novel Protein sim. GBank g135596478 (AF065165) - Myosin-1A [Acanthamoeba castellanii]	UNCLASSIFIED	284488, 29331824, 29331825, 29331826, 285009, 21808754, 264682, 264688, 33657023, 284565

3003	88648079 (6005, 6008)	Novel Protein sim. GBank gji1754989 (U30292) - collagen type XIII alpha-1 chain [Mus musculus]	Contains protein domain (PF01391) - collagen		284512, 284583, 284584, 284587, 284486
3004	88068876 (6007, 6008)	Novel Protein sim. GBank gji2224629(pj)BAAD208021 - (AB002342) KIAA0244 [Homo sapiens]	Contains triple repeat (20 copies)		29331830, 21808789, 284681, 33657109, 283972, 18108385
3005	67794843 (6009, 6010)	Novel Protein sim. GBank gji4880858(pj)AAD27719, 1A-F13294 - (AF13294) CGI-10 protein [Homo sapiens]	Contains protein domain (PF01360) - oxygenase Monooxygenase		29331822, 29331824, 29331827, 60433438, 265011, 265019, 21808766, 21808767, 21808768, 265020, 33657023, 33657349, 60170394, 22279002, 284587
3006	67422224 (6011, 6012)	Novel Protein sim. GBank gji330523 (AF084447) - sex-determination protein homolog Fem1e [Mus musculus]	Contains protein domain (PF00023) - MHC Ank repeat		264259, 28331822, 264512, 21808754, 265018, 284687, 21808765, 284891, 284555, 264556, 284558, 18108385
3007	90938005 (6013, 6014)	Novel Protein sim. GBank gji2385052 (U80738) - CAGH1a [Homo sapiens]	Contains protein domain (PF00096) - Zinc finger, C2H2 type	transcript factor	52844507, 52845156, 65274572, 284909, 264512, 265018, 284760, 284448, 284765, 284889, 60170815, 18108374, 20281152, 284636, 52844332
3008	80416248 (6015, 6016)				264905, 264593, 264766, 264636
3009	81213387 (6017, 6018)	Novel Protein sim. GBank gji3127183 (AF082389) - kidney-specific protein [Rattus norvegicus]	Contains protein domain (PF00501) - AMP-binding enzyme	synthase	52845842, 58182575, 22278995, 22278996, 264259, 29331825, 29331828, 28331827, 29331828, 35698052, 284508, 284509, 264907, 58182435, 264511, 265007, 284512, 285008, 284757, 284758, 55812038, 284759, 33109954, 21808754, 285010, 285011, 284600, 265017, 265018, 265019, 284760, 18108351, 284288, 284388, 21808784, 21808785, 21808787, 55811957, 285020, 265021, 284891, 18108388, 27488282, 20281149, 18108370, 55811576, 284637, 284558, 284557, 18108381, 284558, 58182323, 264559, 18108385, 18108388, 22279002, 284486
3010	85317217 (6019, 6020)	Novel Protein sim. GBank gji4827370(pj)AAD33084, 1A-F06797 - (AF06797) DNA cytosine methyltransferase 3 alpha [Homo sapiens]	Contains protein domain (PF01923) - Protein of unknown function	UNCLASSIFIED	264686, 264887, 21808787, 21808788, 55811957, 22278995, 35695917, 22278996, 22278997, 265020, 285021, 60170815, 264692, 33657023, 28331822, 284693, 18108364, 29331824, 33657109, 60433289, 29331827, 27488281, 28331828, 284508, 284908, 55811576, 35695855, 285008, 284556, 60433438, 83373044, 18108387, 65274727, 60432113, 265017, 22279000, 265019, 284584, 284682, 284764
3011	94333597 (6021, 6022)	Novel Protein sim. GBank gji5052319(pj)AAD38501, 1A-F11883 - (AF11883) citin; edul-onset type II citrullinemia protein [Homo sapiens]	Contains protein domain (PF00153) - Mitochondrial carrier proteins	transport	35696052, 58182435, 284758, 21808754, 265018, 284760, 284762, 18108351, 284682, 284448, 21808766, 65274820, 18108374, 284482, 284564
3012	87753087 (6023, 6024)			UNCLASSIFIED	263972

3013	91236789 (6025, 6026)	Novel Protein sim. GBank gij3702286 (AC005787) - R33374_1 [Homo sapiens]	Contains protein domain (PF00400) - WD domain, G-beta repeat	transcription factor	264468, 263994, 35698288, 22278987, 264239, 29331824, 60424269, 66714117, 35696032, 264905, 264908, 264907, 264908, 264909, 56162435, 264511, 264512, 264870, 264591, 264592, 264593, 264594, 33657402, 60433438, 264595, 264598, 55812038, 264738, 33109934, 21908754, 265010, 265018, 264604, 264780, 264882, 264663, 264764, 264389, 264288, 264785, 264786, 264688, 264768, 264687, 21906767, 35695917, 265020, 33657023, 264692, 264693, 33857109, 264628, 264629, 55811578, 35698423, 35698555, 264630, 264631, 264632, 264634, 264635, 264636, 264637, 264638, 264839, 83373044, 264563, 264565, 264586, 264587 264760
3014	79877263 (6027, 6028)	Novel Protein sim. GBank gij3676374iemh(CAA93087) - (Z88879). Similarly to Yeast Cht12p protein (PIR Acc. No. S54453): cDNA EST EMBL:D27950 comes from this gene; cDNA EST EMBL:D27949 comes from this gene; cDNA EST EMBL:D33447 comes from this gene; cDNA EST EMBL:D33316 comes from ...		ATPase, associated	
3015	86985466 (6029, 6030)				22276995, 22278996, 22278997, 264259, 29331824, 29331828, 264908, 265007, 265008, 264910, 265011, 265017, 265019, 264891, 33657109, 18108370, 35695855, 264558, 264564
3016	87759945 (6031, 6032)	Novel Protein sim. GBank gij1168819spip41733CC91_YEAST - CELL DIVISION CONTROL PROTEIN 91		UNCLASSIFIED	52644507, 52646842, 56994075, 52645080, 29331822, 29331824, 35696052, 33656970, 52644045, 264598, 33657084, 265017, 265019, 52644229, 21906767, 35695917, 52644150, 33657023, 33657109, 27466261, 27466282, 27466284, 33657349, 27466285, 35695783, 35695855, 87168518 264468, 18108387, 22278988, 35696286, 22278989, 264258, 29331822, 60432289, 264908, 29331830, 264909, 56182435, 265008, 265007, 265008, 265009, 264591, 80433356, 60433438, 52648317, 21908754, 55611386, 265010, 265011, 87168559, 265017, 265018, 265019, 264288, 264667, 21906765, 21906766, 21906767, 21908769, 265020, 265022, 85774820, 52645129, 33657108, 33657182, 18108370, 263972, 18108374, 264631, 52644332, 83373044, 18108385, 18108386, 56526486, 87168518, 264404, 60432113, 22279000, 264567
3017	85011154 (6033, 6034)	Novel Protein sim. GBank gij4589658idj(BAA76851.1) - (AB023224) KIAA1007 protein [Homo sapiens]			

3018	11073891 (6035, 6036)	Novel Protein sim. GBank g13218332 (ACD04020) - Unknown gene product [Homo sapiens]		oncogene	264558 264568, 52644507, 18108394, 65274572, 56182575, 22278994, 22278995, 58994075, 22278998, 22278999, 264259, 29331822, 29331824, 60432289, 29331827, 284908, 58182435, 265007, 265009, 60432229, 264593, 60433358, 55812038, 21806754, 87188474, 265011, 87188559, 265017, 265018, 285019, 264881, 18108351, 284448, 264682, 264683, 18108354, 264885, 264887, 264689, 21908788, 21908789, 52644150, 264690, 264691, 33657023, 264692, 284693, 33657109, 52645129, 33657349, 264829, 65274791, 264634, 52644332, 56182323, 18108365, 87188518, 22279000, 22278002, 284583
3020	84318251 (6038, 6040)	Novel Protein sim. GBank g13414809 (AF061529) - r1s [Mus musculus]	Contains protein domain (PF00415) - Regulator of chromosome condensation (RCC1)	ATPase-associated	264488, 283994, 35698286, 264259, 264508, 264905, 264509, 264906, 264907, 264908, 264909, 264510, 264910, 80174839, 264600, 264603, 264760, 264762, 264682, 264763, 264764, 264288, 264389, 264788, 264687, 264688, 284769, 55811957, 35895917, 33657023, 264628, 35898423, 35895855, 264630, 284632, 264834, 264635, 264836, 264637, 264556, 264557, 264638, 83373044, 18108395, 264584, 264587, 264486
3021	60478512 (6041, 6042)	Novel Protein sim. GBank g13808891embJCAB080051 - (Z95559) cDNA EST YK23804.5 comes from this gene; cDNA EST EMBL C13455 comes from this gene; cDNA EST YK32986.5 comes from this gene; cDNA EST CEMSH45R comes from this gene [Caenorhabditis elegans]			264768, 264628, 264482
3022	87716500 (6043, 6044)			UNCLASSIFIED	264259, 29331826, 29331828, 264288, 264568
3023	85305484 (6045, 6046)	Novel Protein sim. GBank g14165825IP232323JAGA1_YEAST - A-AGGLUTININ ATTACHMENT SUBUNIT PRECURSOR	Contains protein domain (PF00814) - Phospholipase D Active site motif	UNCLASSIFIED	264488, 22278995, 35696286, 22278997, 29331826, 35696052, 264807, 29331830, 52644045, 56182435, 60432228, 264592, 60433358, 60433438, 264688, 21908787, 55811957, 35695917, 265021, 18108378, 263978, 264835, 264558, 22279000
3024	86675305 (6047, 6048)			UNCLASSIFIED	60432049, 264760, 21908788, 55811957, 35895917, 264690, 264555, 264559
3025	85706829 (6049, 6050)	Novel Protein sim. GBank g1295671 (L11275) - selected as a weak suppressor of a mutant of the subunit AC40 of DNA dependant RNA polymerase I and III [Saccharomyces cerevisiae]			264593, 55811576

3028	67643862 (6051, 6052)	Novel Protein sim. GBank gij3024052ispjP87924IKAR1_RAT - KALIRIN (PAM COOH-TERMINAL INTERACTOR PROTEIN 10) (P-CIP10)		UNCLASSIFIED	22278996, 22278997, 264400, 29331825, 264111, 265007, 80170831, 265010, 87186559, 265018, 21906765, 29148627, 263987, 20281148, 20281089, 283975, 263977, 20281071, 58528488, 22279000
3027	94844563 (6053, 6054)	Novel Protein sim. GBank gij4928647gijAA034084.1IAF15184 - (AF151847) CGI-89 protein [Homo sapiens]	Contains protein domain (PF01528) - DHC zinc finger domain	UNCLASSIFIED	18106394, 22278995, 22278996, 35696286, 22278997, 22278998, 60432048, 264259, 29331822, 29331824, 66714117, 29331825, 60432288, 29331826, 29331827, 29331828, 35698052, 264905, 264907, 29331830, 264908, 264909, 264510, 265007, 265006, 265009, 264910, 33657402, 264598, 21906754, 265010, 285011, 87188559, 264600, 265016, 18106351, 264682, 264683, 264764, 264288, 264685, 264687, 264769, 264888, 21906765, 21906766, 21906767, 21906768, 21906769, 29148629, 35695917, 265020, 265021, 265022, 52644150, 264692, 33657023, 264893, 52845129, 33657109, 27488281, 18108374, 55811576, 35698423, 65274791, 264638, 264558, 264557, 264638, 60170394, 264639, 264558, 83373044, 18108385, 58528488, 22279000, 22279002
3028	94231897 (6055, 6056)	Novel Protein sim. GBank gij3080521jembCAA18650J - (AL022588) hypothetical protein [Schizosaccharomyces pombe]	Contains protein domain (PF00400) - WD domain, G-beta repeat	UNCLASSIFIED	22278995, 22278996, 22278997, 22278998, 284259, 29331824, 29331827, 35696052, 264908, 265007, 265008, 265009, 60170831, 21906754, 285011, 87168559, 265018, 284782, 264883, 264785, 264889, 21906785, 21906786, 21906789, 29148628, 35695917, 265021, 265022, 33657109, 27488265, 264628, 264629, 18108374, 35698423, 35695855, 264638, 80170394, 22279000, 22279002, 264482, 264564
3029	87618284 (6057, 6058)			UNCLASSIFIED	22278997, 22278998, 29331827, 264905, 284509, 284908, 264510, 264511, 264512, 87168474, 265019, 18108351, 21906768, 264534, 264680, 264883, 263889, 18108370, 264558, 22279000, 22279002, 264482
3030	87544826 (6059, 6060)	Novel Protein sim. GBank gij375726jembCAA187621 - (AL022727) d88019.1 (olfactory receptor-like protein [Hs8M1-1]) [Homo sapiens]	Contains protein domain (PF00001) - 7 transmembrane receptor (rhodopsin family)	tm7	
3031	81877853 (6061, 6062)	Novel Protein sim. GBank gij4530587gijAA022105.11 - (AF132000) TADA1 protein [Homo sapiens]		UNCLASSIFIED	22278995, 22278997, 22278998, 22278999, 284259, 29331822, 29331824, 29331826, 29331827, 29331828, 35698052, 264908, 265018, 21906765, 21906766, 21906767, 21906768, 265021, 263974, 18108374, 264558, 58528488, 22279000, 22279002

3032	94130124 (6083, 6084)	Novel Protein sim. GBank gij1019951 (U37429) - similar to M. musculus MEK5 and other AHC/CTSA proteins [Caenorhabditis elegans]	Contains protein domain (PF00534) - Glycosyl transferases group 1	synthase	22278996, 35696286, 264259, 29331624, 28331626, 264907, 29331830, 264758, 33109954, 8718474, 87168559, 265019, 264288, 21906769, 265021, 264683, 35696423, 35695855, 264638, 58162323, 8373044, 67188516
3033	95308321 (6085, 6086)	Novel Protein sim. GBank gij5031573e[NP_005712.1]ACTR - ARP3 (actin-related protein 3, yeast) homolog	Contains protein domain (PF00022) - Actin	struct	35696286, 264259, 29331826, 35696052, 264508, 264905, 264808, 264907, 264908, 264909, 265008, 264591, 21906754, 285010, 265019, 264681, 264366, 264768, 21906764, 21906768, 35695917, 33657023, 264628, 35695655, 264632, 264635, 264639, 264482, 284563
3034	80415373 (6087, 6088)			UNCLASSIFIED	264906, 264907, 264510, 264592, 285010, 264762, 264766, 264837, 264838, 264486, 264636
3035	91220892 (6089, 6070)	Novel Protein sim. GBank gij3738207[emb]CA212821 - (AL031853) conserved ATP-GTP binding protein [Schizosaccharomyces pombe]		UNCLASSIFIED	264807, 33657402, 285021
3036	91716323 (6071, 6072)	Novel Protein sim. GBank gij728637[sp]P39194[ALU7_HUMAN - III ALU SUBFAMILY SQ WARNING ENTRY III]		kinase	265017
3037	95307434 (6073, 6074)	Novel Protein sim. GBank gij4406590[gb]A020040 - (AF131766) Similar to Era-VASP like protein [Homo sapiens]			
3038	95421807 (6075, 6076)	Novel Protein sim. GBank gij350093[gb]A042865.1[AF15509 - (AF155099) NY-REN 18 antigen [Homo sapiens]	Contains protein domain (PF00627) - UBA domain	UNCLASSIFIED	22278996, 22278997, 264259, 264905, 285007, 265009, 60433356, 21906754, 285016, 285019, 18108351, 264867, 21906785, 285020, 285021, 85274620, 27486282, 264636, 56162323, 16108385, 22279000
3039	87332257 (6077, 6078)	Novel Protein sim. GBank gij4757128[emb]CA842094.1 - (AJ238717) ZRP protein [Rattus norvegicus]		UNCLASSIFIED	35696286, 29331826, 264109, 264110, 264511, 265007, 21908754, 285011, 264681, 264883, 264687, 21906788, 284891, 18106370, 263972, 264629, 18108374, 263977, 35696423, 264594, 18108391, 264892, 264556, 18108382, 18108385, 264567
3040	90833517 (6079, 6080)	Novel Protein sim. GBank gij4884278[emb]CA843247.1 - (AL050037) hypothetical protein [Homo sapiens]			56994075, 22276997, 22276998, 29331827, 33656970, 33108954, 21906754, 87168559, 264800, 264683, 21906765, 21906768, 22279002
3041	88312357 (6081, 6082)	Novel Protein sim. GBank gij3678073[emb]CA804122.1 - (Z61505) similar to Zinc finger, C3HC4 type (RING finger); cDNA EST EMBL:D28025 comes from this gene; cDNA EST EMBL:D28024 comes from this gene; cDNA EST EMBL:D33210 comes from this gene; cDNA EST EMBL:D33441 comes from this ...		UNCLASSIFIED	
3042	85749402 (6083, 6084)	Novel Protein sim. GBank gij790236 (U21156) - sarcolemmal associated protein-2 [Dryopteris curticul]		glycoprotein	264636

3043	87773026 (6085, 8086)	Novel Protein sim. GBank gll654065jemb[CAA58337] - (X63413) U88 [human herpesvirus 6]		UNCLASSIFIED	35686286, 80424269, 35696052, 264508, 264805, 66712502, 56182435, 55811388, 52644296, 55811150, 35695917, 60170615, 33657109, 18108374, 264834, 60431850
3044	87648182 (6087, 6088)	Novel Protein sim. GBank gll4104922 (AF042278) - o251 homolog [Pseudomonas putida]	Contains protein domain (PF01209) - ubiC/COG5 methyltransferase family	glycoprotein	22278988, 22278989, 22278999, 29331824, 56182435, 264511, 265007, 60170831, 60432229, 60433356, 33108954, 18108351, 264288, 35695917, 18108368, 18108370, 60170394
3045	84127598 (6089, 6090)	Novel Protein sim. GBank gll4589860jdbjBAA76659.11 - (AB023232) KIAA1015 protein [Homo sapiens]	Contains protein domain (PF00096) - Zinc finger, C2H2 type	dna_mn_bind	264488, 264259, 35686052, 264508, 264805, 264509, 264908, 264907, 264909, 264511, 265006, 264591, 264593, 33108954, 264604, 264764, 264683, 264288, 294766, 264788, 21806765, 21806766, 55811857, 35895817, 27486282, 18108370, 264828, 18108374, 35695855, 264630, 264632, 264635, 264563, 264564, 264566
3046	88096247 (6091, 6092)			UNCLASSIFIED	22278999, 29331822, 29331824, 29331825, 29331826, 60432289, 29331827, 29331828, 284908, 52846317, 55811957, 60432113, 22279000, 22279002, 264482, 264584
3047	85089824 (6093, 6094)			UNCLASSIFIED	264488, 22278988, 22278989, 22278989, 29331824, 29331825, 56182435, 264511, 265008, 265009, 285011, 285017, 264788, 21906788, 21906789, 35695917, 52644150, 33657349, 85274781, 35695855, 264555, 60432113, 22279000, 264568
3048	87829419 (6095, 6096)	Novel Protein sim. GBank gll4588034jdbjAAD25962.11AF09287 (AF09287) zinc RING finger protein SAG [Homo sapiens]	Contains protein domain (PF00097) - Zinc finger, C2H2 type (RING finger)	UNCLASSIFIED	264102, 28148784
3049	88228955 (6097, 6098)	Novel Protein sim. GBank gll5454158jdbjNF_008286.11pVARS - vally-IRNA synthetase 1	Contains protein domain (PF01408) - RNA synthetases class I (C)	UNCLASSIFIED	22278997, 29331828, 264807, 264758, 87166559, 265018, 264448, 21908768, 285020, 33857109, 35695855, 80432113, 22279000
3050	87643679 (6099, 6100)	Novel Protein sim. GBank gll45898642jdbjBAA76843.11 - (AB023218) KIAA0999 protein [Homo sapiens]	Contains protein domain (PF00089) - Eukaryotic protein kinase domain	kinase	264258, 29331825, 264909, 265007, 264512, 265018, 284288, 21908768, 285020, 264693, 18108365, 58528486, 87168518, 22278989, 264585, 265019, 264288, 264683, 87168518
3051	87750589 (6101, 6102)				264534
3052	57108030 (6103, 6104)	Novel Protein sim. GBank gll117528jdbjP14758jCRYL_FABIT - LAMBDA-CRYSTALLIN		dehydrogenase	

3053	95350373 (6105, 6106)	Novel Protein sim. GBank gij347613jemb CAA19465.1 - (AL023826) cDNA EST EMBL_M89006 comes from this gene; cDNA EST yk282d3.5 comes from this gene [Caenorhabditis elegans]	UNCLASSIFIED	65274572, 56161688, 22276995, 35696286, 22278998, 264259, 60432289, 265008, 285009, 60433438, 21906754, 285010, 67186559, 284603, 265018, 285019, 264763, 284764, 264268, 21906765, 21906766, 21906768, 21906769, 35695917, 18106374, 35696423, 264638, 56182323, 22279000, 264563
3054	86643510 (8107, 6108)	Novel Protein sim. GBank gij1076211jpi S50755 - hypothetical protein VSP-3 - Chlamydomonas reinhardtii	UNCLASSIFIED	35696286, 35696052, 28331830, 264906, 284809, 264512, 264910, 265017, 264604, 264766, 265020, 33657109, 264628, 35695855, 264636, 264564, 264566, 264488
3055	95350337 (6109, 6110)	Novel Protein sim. GBank gij4860855jbp AACD27717.1 AF13294 - (AF132942) CGI:08 protein [Homo sapiens]	transport	80424179, 65274572, 56182575, 35696286, 22278996, 22276999, 60432049, 264259, 60424269, 60432289, 35696052, 56182435, 285008, 265009, 60170631, 60432229, 60431735, 60433356, 264594, 60433438, 21906754, 55811386, 265011, 67168559, 265019, 18108351, 264683, 264288, 264369, 264669, 21908768, 55811857, 35695917, 80170815, 33657023, 65274620, 33657109, 35695763, 60431528, 18106374, 55810764, 55811576, 35696423, 65274791, 264636, 80431850, 16108381, 56182323, 60170394, 16108385, 60432113, 264564, 264565, 264566
3056	91661836 (6111, 6112)	Novel Protein sim. GBank gij728837jpp P39194 ALU7_HUMAN - III ALU SUBFAMILY SQ WARNING ENTRY IIII	glycoprotein	264468, 264589, 18108394, 52846842, 22278997, 22276996, 22276999, 264259, 65714117, 29331826, 29331827, 35696052, 264508, 264509, 264905, 264906, 264907, 264908, 264909, 285006, 264512, 285007, 265008, 265009, 264910, 33857402, 55812036, 284598, 284756, 265010, 265011, 265017, 265019, 264760, 18108351, 264762, 264763, 264764, 264288, 264766, 264687, 16106357, 264768, 264769, 264689, 21906765, 21906766, 21906767, 21906768, 21906769, 35695917, 265020, 285021, 264691, 264693, 33657109, 16106370, 264628, 264829, 16106374, 55811576, 35696423, 35695855, 264630, 264631, 284632, 264634, 264635, 264638, 264637, 264638, 18108381, 83373044, 18108385, 22279000, 22279002, 264563, 264584, 264585, 264586, 264588, 264488, 264567

3057	95412746 (6113, 6114)	Novel Protein sim. GBank gi13878119[emb]CA88601 - (Z48068) similar to GTP-binding protein: cDNA EST EMBL: M89111 comes from this gene: cDNA EST EMBL: D27708 comes from this gene: cDNA EST EMBL: D27708 comes from this gene: cDNA EST EMBL: D73786 comes from this gene: cDNA EST M353..		struct	264508, 264905, 264907, 264908, 264909, 264510, 264512, 264910, 264592, 264594, 264767, 18108374, 264835, 264555, 264837, 264639, 264563, 264564, 264565, 264488
3058	78646226 (8115, 8116)	Novel Protein sim. GBank gi4588034[bl]AAD25882 - (AF09287) - zinc RING finger protein SAG [Homo sapiens]	Contains protein domain (PF00087) - Zinc finger, C2H2 type (RING finger)	UNCLASSIFIED	284893 22278995, 22278996, 22278997, 22278998, 22278999, 264490, 264259, 28331824, 29331825, 29331827, 35886052, 29331828, 265007, 60433438, 265017, 265018, 265019, 264681, 264448, 264288, 264788, 21906785, 21908766, 21908787, 21908789, 29148629, 29148784, 265022, 52644150, 18108370, 264838, 18108385, 264583, 264567, 264587
3059	87829425 (8117, 8118)	Novel Protein sim. GBank gi4588034[bl]AAD25882 - (AF09287) - zinc RING finger protein SAG [Homo sapiens]		UNCLASSIFIED	284112, 52644298, 21908788, 33657023, 283974, 18108385
3060	79346691 (8119, 8120)	Novel Protein sim. GBank gi4454690[bl]AAD208631 - (AF070657) glutathione S-transferase subunit 13 homologue [Homo sapiens]		UNCLASSIFIED	264608, 265008, 18108351, 264566
3061	87740984 (8121, 8122)	Novel Protein sim. GBank gi4454690[bl]AAD208631 - (AF070657) glutathione S-transferase subunit 13 homologue [Homo sapiens]		transferase	18108359, 264558
3062	87619465 (8123, 8124)	Novel Protein sim. GBank gi4454690[bl]AAD208631 - (AF070657) glutathione S-transferase subunit 13 homologue [Homo sapiens]		UNCLASSIFIED	52646365, 52646842, 85274572, 58182575, 56181686, 22278995, 22278998, 22278997, 22278999, 264259, 60433049, 29331824, 66714117, 264508, 264807, 264908, 56182435, 285009, 60432229, 60433438, 55812038, 52844286, 285018, 264682, 264288, 264686, 264768, 264687, 52844229, 264689, 21908768, 264691, 264692, 284883, 18108370, 18108377, 55811576, 264638, 56182323, 264558, 264639, 18108385, 22278000, 22278002
3063	80076023 (8125, 8126)	Novel Protein sim. GBank gi4454690[bl]AAD208631 - (AF070657) glutathione S-transferase subunit 13 homologue [Homo sapiens]	Contains protein domain (PF00403) - Heavy-metal-associated domain	UNCLASSIFIED	22278996, 22278998, 264093, 264094, 264095, 29331824, 60424289, 68714117, 264100, 264807, 265007, 264591, 60432229, 264593, 265011, 265019, 18108351, 264766, 264767, 21908765, 21908788, 264893, 20281069, 22278000, 22278002, 264482, 264568, 264587
3064	91241528 (8127, 8128)	Novel Protein sim. GBank gi4454690[bl]AAD208631 - (AF070657) glutathione S-transferase subunit 13 homologue [Homo sapiens]		UNCLASSIFIED	
3065	91839201 (8128, 8130)	Novel Protein sim. GBank gi4454690[bl]AAD208631 - (AF070657) glutathione S-transferase subunit 13 homologue [Homo sapiens]		UNCLASSIFIED	

3066	91224437 (6131, 6132)	Novel Protein sim. GBank gij4884268jemb CA843245.11 - (AL050028) hypothetical protein [Homo sapiens]		UNCLASSIFIED	16106397, 22278995, 56894075, 22278996, 284905, 68712502, 265008, 284512, 284910, 284758, 60174639, 284760, 18108351, 284764, 284683, 16108359, 284692, 18108354, 16108368, 18106370, 18108377, 18108378, 60170394, 284567
3067	95422551 (6133, 6134)	Novel Protein sim. GBank gij4889258j A027832.1 AF12165 - (AF12165) sorting nexin 9 [Homo sapiens]	Contains protein domain (PF00787) - struct		284486, 284489, 35696286, 22278996, 56994075, 284259, 28331822, 29331625, 35696052, 28331826, 284508, 284905, 284509, 284908, 284907, 284908, 284909, 284112, 284510, 284511, 284512, 285008, 285009, 284910, 284591, 284592, 284593, 284594, 284757, 284595, 284596, 284758, 285010, 285011, 87168559, 284801, 284602, 284603, 284604, 284605, 285019, 284760, 284762, 284448, 284783, 284764, 284288, 284369, 284768, 284786, 284687, 284769, 35695917, 285020, 285021, 284534, 52644150, 284681, 33657023, 284893, 284828, 60431528, 283977, 35695855, 284630, 284631, 284634, 284635, 284638, 284837, 284638, 284639, 63373044, 56526486, 87168518, 22278000, 22278002, 284563, 284483, 284584, 284565, 284566, 284567, 284488
3068	85360851 (6135, 6136)	Novel Protein sim. GBank gij3678119jemb CAA88601 - (Z49086) similar to GTP-binding protein; cDNA EST	Contains protein domain (PF01826) - struct		284112, 22278996, 56994075, 22278998, 22278999, 284259, 284107, 284905, 28331630, 52644045, 284110, 60170631, 284592, 284594, 33657402, 21906754, 33109954, 87168474, 87168559, 285017, 284446, 284764, 284683, 284768, 52644228, 21906765, 21906786, 21906788, 21906789, 60170615, 33657023, 16108370, 18108376, 284634, 284557, 60170394, 56182323, 18108385, 87168516, 22279000, 284482
3069	95412753 (6137, 6138)	Novel Protein sim. GBank gij3678119jemb CAA88601 - (Z49086) similar to GTP-binding protein; cDNA EST	GTPase of unknown function		

3070	94319173 (6139, 6140)	Novel Protein sim. GBank gll307766[emb]CAB055271 - (Z63110) cDNA EST yk472b5.3 comes from this gene: cDNA EST yk474b7.3 comes from this gene: cDNA EST yk472b5.5 comes from this gene: cDNA EST yk486c10.3 comes from this gene: cDNA EST yk486c10.5 comes from this gene: cDNA EST EM...		synthase	264488, 22278994, 22278995, 22278996, 56994075, 22278997, 22278999, 264259, 29331822, 29147620, 29331824, 66714117, 29331828, 29146496, 29146499, 66712502, 29331830, 52644045, 56182435, 264511, 265007, 264512, 264910, 60170831, 264592, 264756, 33109954, 21906754, 67108474, 265019, 18106351, 264448, 264683, 264288, 52644229, 264689, 21906765, 21906766, 21906787, 21906789, 35895917, 265020, 265021, 60170815, 52644150, 264691, 33657023, 27486261, 27486264, 264628, 18108370, 18108377, 55811578, 35695655, 284634, 264635, 18108381, 60170394, 56182323, 264558, 83373044, 18108385, 18108387, 56526486, 264404, 264563, 264566
3071	94325573 (6141, 6142)	Novel Protein sim. GBank gll4502425[en]NP_001709.1[BMF6 - bone morphogenetic protein 6 precursor	Contains protein domain (PF00085) - bgl		264488, 65274572, 18108396, 22278996, 35696286, 22278997, 22278998, 22278999, 264259, 29331822, 66714117, 29331826, 35696052, 29331826, 29146498, 29146499, 264907, 264908, 29331830, 264909, 52644045, 56182435, 265008, 265007, 264512, 265008, 265009, 60170831, 60432229, 264592, 60433356, 33657402, 60433436, 33109954, 52644286, 87168474, 265010, 265017, 264681, 264288, 264685, 264766, 264687, 264769, 264689, 21908765, 21908766, 21908767, 21908768, 21908769, 35695917, 265020, 265021, 265022, 80170615, 52644150, 264690, 264691, 264692, 33657023, 264683, 33657109, 263971, 18108374, 18108377, 35696423, 55811576, 65274791, 35695855, 264630, 264635, 264636, 264557, 60170394, 83373044, 60432113, 22279000, 22279002, 264563, 264564, 264565, 264566, 264567, 56182375, 29331822, 29331824, 29331825, 29146498, 264908, 52644045, 56182435, 265009, 60433438, 55812038, 18108351, 264683, 264369, 52644229, 52644150, 33657023, 264683, 33657108, 18108374, 55811576, 65274791, 264555, 56182323, 60432113, 264564
3072	95115692 (6143, 6144)	Novel Protein sim. GBank gll1283289 (U47856) - fibrin-4 (Acanus diadematus)		transcriptfactor	

3073	88147246 (6145, 6146)	Novel Protein sim. GBank gij1348405p22528(CORR_HUMAN - CORNIFIN B (SMALL PROLINE-RICH PROTEIN IB) (SPR-IB) (14.9 KD PANCRORNLIN)	UNCLASSIFIED	284789	18108398, 29331822, 29331827, 60432228, 265017, 264691, 264693
3074	880893351 (6147, 6148)	Novel Protein sim. GBank gij3419847 (AC004882) - similar to yeast hypothetical protein ybk4; similar to P38164 (P1D:G588463) [Homo sapiens]	UNCLASSIFIED	264488, 265018, 264448, 264288, 21906767, 264693, 18108388, 18108370, 18108374, 264567	85274572, 35898052, 284511, 60170831, 87168474, 264388, 35895917, 33657182, 27488264, 33657348, 35895783, 35895855, 264639
3075	88085752 (6149, 6150)	Novel Protein sim. GBank gij4557349(efNP_000458, 11BARD - BRCA1 associated RING domain 1	homeobox	264508, 264907, 264688, 264693, 56526486	
3076	87819219 (6151, 6152)		UNCLASSIFIED	18108398, 29331822, 29331827, 60432228, 265017, 264691, 264693	
3077	88734277 (6153, 6154)	Novel Protein sim. GBank gij3023956ip00808(HET1_PODAN - VEGETATIBLE INCOMPATIBILITY PROTEIN HET-E-1	Kinase WD domain, G-beta repeat	265017, 264691, 264693	85274572, 35898052, 284511, 60170831, 87168474, 264388, 35895917, 33657182, 27488264, 33657348, 35895783, 35895855, 264639
3078	880893355 (6155, 6156)	Novel Protein sim. GBank gij3900850 (AC004994) - similar to KIAA0600; similar to d1028456 (P1D:G3043724) [Homo sapiens]		22279002	
3079	87821893 (6157, 6158)	Novel Protein sim. GBank gij3875410(emb(CA602876) - (Z81052) Similarly to Yeast ABC1P protein (SW/ABC1_YEAST): cDNA EST yk2298.3 comes from this gene: cDNA EST yk2298.5 comes from this gene [Caenorhabditis elegans]	transport	29331824, 29331826, 264758, 55811386, 265017, 55811150, 52844228, 21906768, 285020, 265021, 264693, 18108376, 264631, 52844332, 22279002	
3080	95296274 (6159, 6160)	Novel Protein sim. GBank gij5237221(gbAADA1265-11 - (AF117887) protein arginine methyltransferase [Mus musculus]	interferon	264488, 52644507, 22278996, 22278998, 264490, 264259, 29331824, 88714117, 29331825, 29331826, 29331827, 29331828, 29146489, 264508, 264805, 264826, 52644045, 56182435, 265008, 264591, 264596, 21906754, 80174639, 285010, 264682, 264448, 264783, 264683, 264764, 264288, 264685, 264788, 264688, 264689, 21906765, 21906787, 21908769, 55811957, 35895917, 265020, 60170815, 52844150, 264892, 33657023, 264693, 85274820, 33857109, 27488281, 35895783, 264688, 18108370, 65274781, 264558, 56182323, 60170394, 284482, 264585, 264484	
3081	88094864 (6161, 6162)	Novel Protein sim. GBank gij28831sepP39168(ALU1_HUMAN - III ALU SUBFAMILY J WARNING ENTRY IIII	UNCLASSIFIED	18108398, 264509, 264805, 264806, 264907, 264908, 264909, 264510, 264511, 265009, 264910, 264595, 264758, 265011, 285016, 264760, 264781, 264763, 264784, 18108354, 264685, 264766, 264628, 264629, 264630, 264631, 264632, 264634, 264635, 264555, 264636, 18108382, 18108385, 264563, 264565, 264568	264784, 55811957, 264555, 264564
3082	80310121 (6163, 6164)				

3083	86095726 (6165, 6168)	Novel Protein sim. GBank gij663241 (U29488) - C56C10.3 gene product [Caenorhabditis elegans]	UNCLASSIFIED	264488, 264259, 29331824, 264106, 265008, 264591, 264592, 21908754, 264288, 264787, 21906766, 21906769, 29140794, 264691, 264632, 22278000
3084	87445566 (6167, 6168)	Novel Protein sim. GBank gij47674 [gij47674] - probable structural component p38 - borna disease virus		22278995, 80432289, 35696052, 264905, 264906, 264907, 264908, 264909, 265006, 265007, 264910, 264593, 264595, 264758, 264369, 264288, 264766, 35693917, 265020, 18108374, 35696423, 294631, 264556, 264565, 264566, 264567, 264486, 265011, 264681
3085	87795781 (6169, 6170)	Novel Protein sim. GBank gij2565057 (U80741) - CAGH44 [Homo sapiens]	UNCLASSIFIED	22278998, 264092, 264259, 29331822, 29331825, 264108, 264112, 18108351, 264687, 263967, 263974, 55810764, 263981, 18108385, 264467
3086	87769942 (6171, 6172)	Novel Protein sim. GBank gij3894189 (AC005662) - hypothetical protein [Arabidopsis thaliana]	UNCLASSIFIED	52846385, 56894075, 22278997, 22278998, 29331824, 29331825, 35696052, 60433438, 33109954, 21906754, 52846317, 285017, 264862, 264369, 264884, 21906767, 21906768, 265020, 264691, 33657023, 33657109, 52845129, 33857162, 27488262, 35695855, 67168518
3087	87462886 (6173, 6174)			264591
3088	81224441 (6175, 6176)	Novel Protein sim. GBank gij3353304 (AF001549) - Unknown gene product [Homo sapiens]	UNCLASSIFIED	16106397, 65274572, 56182575, 56181668, 56894075, 35696288, 22278997, 22278998, 264259, 29331824, 29331825, 29331826, 29331828, 284907, 29331630, 284809, 56182435, 264510, 265007, 80170831, 60432228, 21906754, 55811366, 265017, 265018, 265019, 264760, 55811150, 264288, 264766, 56181582, 21908785, 21906766, 21906767, 21906768, 265021, 60170815, 27486262, 18108370, 60431528, 35696423, 264558, 264559, 60433213, 264488
3089	85361242 (6177, 6178)	Novel Protein sim. GBank gij4689146 [gij4689146] - (AF077049) lambda-crystallin [Homo sapiens]	Contains protein domain (PF00725) - 3-hydroxyacyl-CoA dehydrogenase	dehydrogenase

3090	95342371 (6178, 6180)	Novel Protein sim. GBank g11354050 (U47024) - MEM3 [Mus musculus]	UNCLASSIFIED	60424178, 52645158, 85274572, 56182575, 56181686, 22278985, 35698288, 56994075, 22278998, 22278998, 22278998, 264259, 28331822, 56162181, 28331824, 28331825, 28331828, 29331827, 28331828, 35698052, 33656970, 264908, 264908, 52644045, 264828, 265006, 265007, 265008, 60170831, 60432229, 60433358, 33657402, 55812038, 264758, 21808754, 33109854, 52646317, 55811388, 52644286, 87188474, 265011, 87166559, 265017, 265018, 265019, 55811150, 18108351, 264861, 264448, 264288, 264369, 18108357, 264768, 52644228, 56181582, 21908764, 21908765, 21908768, 21908767, 21908768, 21908769, 35695917, 265020, 265022, 60170815, 35695917, 265020, 264691, 33857023, 264890, 52644150, 264691, 33857023, 18108365, 85274820, 33857109, 18108368, 33657182, 27486281, 27486285, 35695763, 18108374, 18108378, 55810784, 35696423, 55811578, 65274791, 35695855, 264557, 56182323, 83373044, 18108387, 18108388, 87168518, 22278000, 22278002, 264583, 264482
3091	95317424 (6161, 6162)	Novel Protein sim. GBank g138739321emb1 (CA801859) - (Z79596) Similarity to Bovine aspartyl beta hydroxylase (TR:G162894); cDNA EST EMBL:D27916 comes from this gene; cDNA EST EMBL:D27915 comes from this gene; cDNA EST EMBL:D64881 comes from this gene; cDNA EST EMBL:D68139 comes f...	UNCLASSIFIED	35696288, 29331822, 35698052, 264508, 264509, 264805, 264906, 264908, 264908, 264510, 264758, 265010, 265011, 264883, 264885, 264788, 264788, 264788, 264789, 264893, 264828, 35696423, 35695855, 264632, 264635, 264639, 264482, 264583, 264486

3092	95314592 (6183, 6184)	Novel Protein sim. GBank g117107561pP15680FS2, HUMAN - 40S RIBOSOMAL PROTEIN S2 (S4) (LUREP3 PROTEIN)	Contains protein domain (PF00333) - ribosomal prot Ribosomal protein S5		264468, 60424179, 16108396, 22278985, 56994075, 22278998, 35698288, 22278997, 22278996, 60432049, 264259, 29331622, 29331624, 29331825, 29331826, 29331627, 35898052, 29331828, 29146498, 29146499, 264508, 264509, 264905, 264906, 264907, 29331830, 264908, 264909, 264113, 264510, 264511, 265008, 264512, 265007, 265006, 264910, 265009, 60170631, 264591, 264592, 60431735, 264583, 264594, 60433438, 264595, 264756, 21908754, 265010, 265011, 264601, 264602, 265017, 264603, 264604, 265018, 264605, 265019, 264760, 264762, 264661, 16108351, 264763, 264662, 264446, 264764, 264663, 264288, 264369, 264765, 264766, 264668, 264767, 264667, 264766, 264769, 264669, 21908764, 264669, 21908765, 21908766, 21908767, 21908768, 21908769, 29146829, 29148764, 35895917, 265020, 265021, 264534, 60170615, 264690, 264691, 264692, 65274620, 33657109, 27466262, 264628, 264629, 16108374, 263976, 16108377, 35698423, 264630, 264631, 264632, 264634, 264635, 264555, 264636, 264637, 264556, 264638, 264557, 264558, 264639, 60170394, 16108385, 264259, 28331624, 35698052, 264905, 265008, 60432229, 60431735, 264684, 264369, 264288, 264768, 21908767, 35698423, 63373044, 16108385
3093	94318457 (6185, 6186)	Novel Protein sim. GBank g15002567embjCAB44347.11 - (Y17454) LSFR1 protein [Homo sapiens]		UNCLASSIFIED	
3094	94316675 (6187, 6188)	Novel Protein sim. GBank g14007341pP310441BP, RAT - PHOSPHATIDYLETHANOLAMINE-BINDING PROTEIN (23 KD MORPHINE-BINDING PROTEIN) (P23K)	Contains protein domain (PF01161) - collagen Phosphatidylethanolamine-binding protein		16108398, 264259, 60432289, 29331827, 264511, 264783, 264286, 264787, 265022, 264691, 264693, 65274791, 56182323, 264564, 264565

3095	84646162 (6189, 6190)	Novel Protein sim. GBank gik67759jgblAAC31421.1(A-F12444 - (A-F12440) MAGE tumor antigen D1 [Homo sapiens])	MAGE family	Contains protein domain (PF01454) - UNCLASSIFIED	18106397, 56182575, 22276995, 35666266, 56994075, 22278997, 22278999, 284259, 60432049, 66714117, 29331825, 60432269, 35666052, 33656970, 29146499, 284506, 284905, 284509, 29331830, 284909, 284510, 284511, 284512, 285007, 285008, 285009, 80170831, 284758, 21906754, 65658542, 265010, 285011, 87168559, 265017, 285018, 265019, 284760, 284681, 284882, 284683, 284764, 284369, 284288, 284886, 284766, 284769, 284689, 21906765, 21908768, 21908767, 55811957, 35895917, 285020, 285021, 285022, 52644150, 284691, 284692, 33657023, 284693, 283972, 18108378, 55811578, 35898423, 284952, 60170394, 284639, 63373044, 18108365, 18108387, 65274727, 87168518, 60432113, 284482, 284563, 284564, 284565, 284487, 18108391, 22278995, 22278996, 22278997, 22278999, 28331824, 29331825, 29331826, 29331827, 33656970, 284905, 284908, 285006, 284910, 33657402, 285011, 285017, 285018, 284369, 21908766, 21908767, 21908768, 35895917, 285020, 60170615, 284691, 284692, 284693, 27486261, 27486262, 18108370, 60431528, 284634, 284635, 284636, 22279000, 284358, 284488, 29331822, 29331825, 60432269, 29331826, 35666052, 29331828, 29331830, 284594, 55812038, 33109954, 33657084, 87168474, 87168559, 52844229, 21908765, 21908767, 18108376, 35898423, 52844332, 284636, 60432113, 22279002, 284634, 284637, 284565, 284486, 35666266, 29331822, 29331824, 29331825, 29331827, 285007, 285008, 285010, 285011, 285016, 285018, 18109357, 21908768, 285020, 285022, 55811576, 56182323, 22279002, 284563, 29331825, 60432269, 35666052, 284910, 60432229, 284592, 284288, 284693, 283967, 284635, 284908, 284693, 284628, 284630, 284632
3096	67756128 (6191, 6192)	Novel Protein sim. GBank gik3862221[dbj]BAJ34470.1 - (AB018293) KIAA0750 protein [Homo sapiens]	Contains protein domain (PF00307) - struct	UNCLASSIFIED	284563, 284564, 284565, 284487, 18108391, 22278995, 22278996, 22278997, 22278999, 28331824, 29331825, 29331826, 29331827, 33656970, 284905, 284908, 285006, 284910, 33657402, 285011, 285017, 285018, 284369, 21908766, 21908767, 21908768, 35895917, 285020, 60170615, 284691, 284692, 284693, 27486261, 27486262, 18108370, 60431528, 284634, 284635, 284636, 22279000, 284358, 284488, 29331822, 29331825, 60432269, 29331826, 35666052, 29331828, 29331830, 284594, 55812038, 33109954, 33657084, 87168474, 87168559, 52844229, 21908765, 21908767, 18108376, 35898423, 52844332, 284636, 60432113, 22279002, 284634, 284637, 284565, 284486, 35666266, 29331822, 29331824, 29331825, 29331827, 285007, 285008, 285010, 285011, 285016, 285018, 18109357, 21908768, 285020, 285022, 55811576, 56182323, 22279002, 284563, 29331825, 60432269, 35666052, 284910, 60432229, 284592, 284288, 284693, 283967, 284635, 284908, 284693, 284628, 284630, 284632
3097	88284895 (6193, 6194)	Novel Protein sim. GBank gik468288[emb]CA8379611 - (AL022395) d4273N12.1 (PUTATIVE protein based on EST matches) [Homo sapiens]	Contains protein domain (PF00646) - UNCLASSIFIED	UNCLASSIFIED	284563, 284564, 284565, 284487, 18108391, 22278995, 22278996, 22278997, 22278999, 28331824, 29331825, 29331826, 29331827, 33656970, 284905, 284908, 285006, 284910, 33657402, 285011, 285017, 285018, 284369, 21908766, 21908767, 21908768, 35895917, 285020, 60170615, 284691, 284692, 284693, 27486261, 27486262, 18108370, 60431528, 284634, 284635, 284636, 22279000, 284358, 284488, 29331822, 29331825, 60432269, 29331826, 35666052, 29331828, 29331830, 284594, 55812038, 33109954, 33657084, 87168474, 87168559, 52844229, 21908765, 21908767, 18108376, 35898423, 52844332, 284636, 60432113, 22279002, 284634, 284637, 284565, 284486, 35666266, 29331822, 29331824, 29331825, 29331827, 285007, 285008, 285010, 285011, 285016, 285018, 18109357, 21908768, 285020, 285022, 55811576, 56182323, 22279002, 284563, 29331825, 60432269, 35666052, 284910, 60432229, 284592, 284288, 284693, 283967, 284635, 284908, 284693, 284628, 284630, 284632
3098	80258024 (6195, 6196)	Novel Protein sim. GBank gik303603[dbj]BA02145.1 - (D12821) cytochrome P-4501BV [Homo sapiens]	cyto450	UNCLASSIFIED	284563, 284564, 284565, 284487, 18108391, 22278995, 22278996, 22278997, 22278999, 28331824, 29331825, 29331826, 29331827, 33656970, 284905, 284908, 285006, 284910, 33657402, 285011, 285017, 285018, 284369, 21908766, 21908767, 21908768, 35895917, 285020, 60170615, 284691, 284692, 284693, 27486261, 27486262, 18108370, 60431528, 284634, 284635, 284636, 22279000, 284358, 284488, 29331822, 29331825, 60432269, 29331826, 35666052, 29331828, 29331830, 284594, 55812038, 33109954, 33657084, 87168474, 87168559, 52844229, 21908765, 21908767, 18108376, 35898423, 52844332, 284636, 60432113, 22279002, 284634, 284637, 284565, 284486, 35666266, 29331822, 29331824, 29331825, 29331827, 285007, 285008, 285010, 285011, 285016, 285018, 18109357, 21908768, 285020, 285022, 55811576, 56182323, 22279002, 284563, 29331825, 60432269, 35666052, 284910, 60432229, 284592, 284288, 284693, 283967, 284635, 284908, 284693, 284628, 284630, 284632
3099	91243325 (6197, 6198)	Novel Protein sim. GBank gik303603[dbj]BA02145.1 - (D12821) cytochrome P-4501BV [Homo sapiens]	cyto450	UNCLASSIFIED	284563, 284564, 284565, 284487, 18108391, 22278995, 22278996, 22278997, 22278999, 28331824, 29331825, 29331826, 29331827, 33656970, 284905, 284908, 285006, 284910, 33657402, 285011, 285017, 285018, 284369, 21908766, 21908767, 21908768, 35895917, 285020, 60170615, 284691, 284692, 284693, 27486261, 27486262, 18108370, 60431528, 284634, 284635, 284636, 22279000, 284358, 284488, 29331822, 29331825, 60432269, 29331826, 35666052, 29331828, 29331830, 284594, 55812038, 33109954, 33657084, 87168474, 87168559, 52844229, 21908765, 21908767, 18108376, 35898423, 52844332, 284636, 60432113, 22279002, 284634, 284637, 284565, 284486, 35666266, 29331822, 29331824, 29331825, 29331827, 285007, 285008, 285010, 285011, 285016, 285018, 18109357, 21908768, 285020, 285022, 55811576, 56182323, 22279002, 284563, 29331825, 60432269, 35666052, 284910, 60432229, 284592, 284288, 284693, 283967, 284635, 284908, 284693, 284628, 284630, 284632
3100	87602421 (6199, 6200)	Novel Protein sim. GBank gik108376[dbj]BA013 - rRNA protein precursor, parolite - rat	UNCLASSIFIED	UNCLASSIFIED	284563, 284564, 284565, 284487, 18108391, 22278995, 22278996, 22278997, 22278999, 28331824, 29331825, 29331826, 29331827, 33656970, 284905, 284908, 285006, 284910, 33657402, 285011, 285017, 285018, 284369, 21908766, 21908767, 21908768, 35895917, 285020, 60170615, 284691, 284692, 284693, 27486261, 27486262, 18108370, 60431528, 284634, 284635, 284636, 22279000, 284358, 284488, 29331822, 29331825, 60432269, 29331826, 35666052, 29331828, 29331830, 284594, 55812038, 33109954, 33657084, 87168474, 87168559, 52844229, 21908765, 21908767, 18108376, 35898423, 52844332, 284636, 60432113, 22279002, 284634, 284637, 284565, 284486, 35666266, 29331822, 29331824, 29331825, 29331827, 285007, 285008, 285010, 285011, 285016, 285018, 18109357, 21908768, 285020, 285022, 55811576, 56182323, 22279002, 284563, 29331825, 60432269, 35666052, 284910, 60432229, 284592, 284288, 284693, 283967, 284635, 284908, 284693, 284628, 284630, 284632
3101	79602134 (6201, 6202)	UNCLASSIFIED	UNCLASSIFIED	UNCLASSIFIED	284563, 284564, 284565, 284487, 18108391, 22278995, 22278996, 22278997, 22278999, 28331824, 29331825, 29331826, 29331827, 33656970, 284905, 284908, 285006, 284910, 33657402, 285011, 285017, 285018, 284369, 21908766, 21908767, 21908768, 35895917, 285020, 60170615, 284691, 284692, 284693, 27486261, 27486262, 18108370, 60431528, 284634, 284635, 284636, 22279000, 284358, 284488, 29331822, 29331825, 60432269, 29331826, 35666052, 29331828, 29331830, 284594, 55812038, 33109954, 33657084, 87168474, 87168559, 52844229, 21908765, 21908767, 18108376, 35898423, 52844332, 284636, 60432113, 22279002, 284634, 284637, 284565, 284486, 35666266, 29331822, 29331824, 29331825, 29331827, 285007, 285008, 285010, 285011, 285016, 285018, 18109357, 21908768, 285020, 285022, 55811576, 56182323, 22279002, 284563, 29331825, 60432269, 35666052, 284910, 60432229, 284592, 284288, 284693, 283967, 284635, 284908, 284693, 284628, 284630, 284632

3102	91220892 (8203, 8204)	Novel Protein sim. GBank gll5305708lp/AAAD1781.1(AE128536) cytoplasmic phosphoprotein PACSIN2 (Homo sapiens)	Contains protein domain (PF00016) - SH3 domain	struct	35696286, 22278996, 22278999, 29331827, 35696052, 264808, 264512, 265008, 60170831, 60433358, 33109954, 18108351, 264884, 264689, 21906767, 60170615, 264692, 33657023, 264638, 22279000, 264482, 264584 35695917, 264565
3103	90838004 (8205, 8206)	Novel Protein sim. GBank gll464584lp/P35282/RB17_MOUSE - RAS-RELATED PROTEIN RAB-17		UNCLASSIFIED	
3104	87340833 (8207, 8208)	Novel Protein sim. GBank gll5032207/etfjNP_005898.1pTSSC - tumor-suppressing STF cDNA 6		UNCLASSIFIED	264259, 264684, 264532, 33657182, 264558
3105	84148603 (8209, 8210)				22278997, 264259, 29331824, 35686052, 29331826, 264508, 264509, 264905, 264908, 264907, 264908, 264511, 264910, 264591, 264594, 264758, 264760, 264681, 264762, 264784, 264288, 264788, 264788, 264887, 284789, 21906768, 21906788, 35695917, 33657023, 264692, 264693, 264628, 264829, 35695855, 264830, 264631, 264632, 264834, 264835, 264837, 264838, 264839, 83373044, 264404, 22278002, 264563, 264565, 264568, 264486, 264587
3108	85361416 (8211, 8212)	Novel Protein sim. GBank gll1938574 (U97190) - B0025.2 gene product [Caenorhabditis elegans]			22278996, 22278997, 22278998, 22278999, 264092, 264093, 264094, 29331822, 264908, 264907, 264908, 52844045, 58182435, 264112, 265008, 265009, 55612038, 265017, 285018, 264883, 264686, 264687, 264788, 52844229, 21906765, 21906768, 21906789, 55811957, 265020, 265022, 264690, 52844150, 264692, 264693, 18108370, 18108377, 55611578, 56182323, 18108385, 18108388, 22278000, 264563
3107	85343272 (8213, 8214)	Novel Protein sim. GBank gll3341441/emb(CAA76851) - (Y17794) winged-helix transcription factor [Gallus gallus]			22278995, 22278996, 35696288, 22278997, 22278998, 264091, 264093, 264259, 29331822, 29331825, 29331826, 60432289, 29331827, 29331828, 33656970, 264105, 264512, 265009, 60433358, 60433438, 265011, 265017, 265018, 21906765, 21906768, 21906789, 265021, 21906789, 21906789, 265021, 264891, 33657109, 27488261, 27488265, 18108370, 263972, 18108374, 55611578, 18108385, 55626488, 264482, 284487
3108	87340835 (8215, 8216)	Novel Protein sim. GBank gll5032207/etfjNP_005898.1pTSSC - tumor-suppressing STF cDNA 8		UNCLASSIFIED	56182435, 264288, 264690, 264564

3109	94318461 (6217, 6216)	Novel Protein sim. GBank gij5002567[emb]CAB44347.1 - (Y17454) LSFR1 protein [Homo sapiens]	Contains protein domain (PF00096) - Zinc finger, C2H2 type	struct	264480, 264608, 285007, 264810, 264593, 264683, 264684, 264687, 21906767, 21906780, 264693, 18108370, 264629, 18108374, 264632, 264638, 22279000
3110	95090716 (6218, 6220)	Novel Protein sim. GBank gij10762110[emb]550755 - hypothetical protein VSP-3 - Chlamydomonas reinhardtii		UNCLASSIFIED	264684, 65274572, 22278995, 22278997, 26432049, 264259, 29331822, 29331824, 26331825, 60432289, 29331826, 29331827, 29331828, 264608, 264510, 265008, 265007, 265008, 265009, 60432228, 33857402, 60433356, 265011, 87168559, 264600, 265017, 265018, 265019, 18108351, 264288, 264369, 21906768, 21906787, 21908788, 265020, 60170615, 264693, 65274620, 18108370, 264639, 18108384, 22279000, 264563, 18108390
3111	87754512 (6221, 6222)	Novel Protein sim. GBank gij328231 (U75454) - C2H2 type zinc finger protein [Homo sapiens]	Contains protein domain (PF00096) - Zinc finger, C2H2 type	transcriptflator	264488, 18108398, 68712502, 285017, 285018, 285019, 284448, 21908787, 285020, 33857023, 18108385, 18108388, 35898423, 52644332, 18108385, 18108388
3112	88043839 (6223, 6224)	Novel Protein sim. GBank gij3900848 (AC005023) - match to EST AA381117 (NM_02013438) [Homo sapiens]	Contains protein domain (PF00046) - homeobox domain	homeobox	18108387, 22278989, 264258, 29331824, 358986052, 264607, 264757, 60433438, 87168559, 264783, 264448, 18108354, 264288, 21906787, 21906788, 35895917, 264690, 264691, 264692, 264693, 18108365, 18108381, 18108384, 18108385, 18108388, 87168518, 22279000, 22279002
3113	88207088 (6225, 6226)	Novel Protein sim. GBank gij2459910 (AF005856) - anon245 [Drosophila yakuba]		hm7	284909, 58182435, 264910, 21906754
3114	79843167 (6227, 6228)	Novel Protein sim. GBank gij4988270[emb]AAB52261.2 - (U87002) similar to acyl-CoA dehydrogenases and epoxide hydrolases; Pfam domain PF00441 (Acyl-CoA_dh), Score=57.4, E-value=1.7e-16, N=2; contains similarity to Pfam domain PF00702 (Hydrolase), Score=57.4, E-value=1e-13, N=1	Contains protein domain (PF00702) - hydrolase	hydrolase	26424179, 58182575, 264259, 29331824, 60424269, 29331828, 68712502, 264510, 285007, 80431735, 80433358, 55812038, 55811386, 285019, 264288, 264689, 21906769, 264691, 33857023, 264693, 60431528, 283974, 80431650, 58182323, 264559, 22279000, 22279002
3115	94117896 (6229, 6230)	Novel Protein sim. GBank gij5032225[emb]NP_005678.1pWBS - Williams Beuren syndrome chromosome region 11		transcriptflator	264905, 264758, 21906764, 264690
3116	79842855 (6231, 6232)			UNCLASSIFIED	284510, 265011, 18108351, 264288, 264689, 284691, 18108388, 18108372, 283981, 284558, 264564
3117	87771288 (6233, 6234)			UNCLASSIFIED	

3118	84665848 (6235, 6236)	Novel Protein sim. GBank gij380563jemb CAB01444.11 - (Z78018) predicted using GeneRinder, similar to serine/threonine kinase; cDNA EST yk353d1.0.5 comes from this gene [Caenorhabditis elegans]	Contains protein domain (PF00008) - Igr		52645156, 52646842, 65274572, 56162575, 22276995, 58994075, 22276998, 35686286, 22276997, 22276998, 22276999, 264259, 29331822, 29331824, 66714117, 29331828, 29331827, 35896052, 29331828, 264905, 264908, 29331830, 52644045, 56182435, 264510, 264511, 265007, 285008, 285009, 264757, 52648317, 21908754, 33657084, 52644286, 87168474, 67168559, 265017, 265018, 264805, 265019, 264782, 264448, 264882, 264884, 264288, 264786, 56181562, 21908765, 21908766, 21906768, 21906769, 265020, 265022, 264690, 52644150, 264891, 33657023, 264693, 33657109, 33657349, 264628, 18108370, 60431528, 18108374, 35698423, 65274791, 60170394, 83373044, 87168518, 22279000, 22279002, 264488, 265008, 264288
3119	85720786 (6237, 6238)		Contains protein domain (PF00328) - Histidine acid phosphatase		264488, 264509, 264510, 264511, 264512, 264288, 264486
3120	67344040 (6239, 6240)	Novel Protein sim. GBank gij5019819gij AAC37883.1 AF14315 - (AF143152) putative NADH oxidoreductase complex I subunit [Caenorhabditis elegans]	UNCLASSIFIED		52644507, 52645158, 52646365, 52646842, 22278994, 56994075, 22276996, 22278999, 264259, 29331824, 29331827, 35896052, 52644045, 265008, 52646317, 87168474, 87168559, 21906765, 52644150, 33657023, 18108374, 264637
3121	84110735 (6241, 6242)	Novel Protein sim. GBank gij4501877 ref NP_001088.1 PACRI - acrosin			18108374, 264637
3122	11614528 (6243, 6244)	Novel Protein sim. GBank gij2439517 (AC002563) - putative RHORAC effector protein, 95% similarity to P49205 (P1D:G1345680) [Homo sapiens]	Contains protein domain (PF00780) - CNH domain	UNCLASSIFIED	264638
3123	86063003 (6245, 6246)			kinase	18108382, 29331822, 29331824, 29331825, 284905, 265007, 55812038, 265019, 18108351, 284682, 264286, 264768, 21908764, 21908765, 21908766, 21908769, 55811957, 18108365, 16108368, 27486285, 18108374, 18108381, 18108384, 22279000, 22279002, 264482
3124	87788898 (6247, 6248)			UNCLASSIFIED	264905
3125	81216807 (6249, 6250)	Novel Protein sim. GBank gij4980628gij AAD35412.1 AE00171 - (AE001714) oxidoreductase, short chain dehydrogenase/reductase family [Thermotoga maritima]	Contains protein domain (PF00106) - short chain dehydrogenase	dehydrogenase	56181666, 264259, 66714117, 60432288, 29331826, 29331827, 264807, 284808, 264828, 285009, 60433356, 33657402, 60433438, 264758, 16108351, 264288, 29148627, 29148628, 33657023, 33657109, 18108382, 56526488

3128	95337205 (6251, 6252)			UNCLASSIFIED	22276989, 264480, 264259, 60432049, 29331622, 60432289, 29148498, 52644045, 58182435, 265009, 60433438, 265010, 87168559, 265017, 265018, 55811150, 264783, 264883, 264389, 264685, 29146629, 33657023, 264693, 33657109, 18108374, 55811576, 18108365, 60432113, 22279002, 33659286, 22276998, 22276999, 28331828, 264908, 60433438, 87168559, 264604, 21906765, 21908769, 33657023, 33657349, 264629, 18108374, 18108377, 22279000, 22279002
3127	91639233 (6253, 6254)	Novel Protein sim. GBank g12828280jemb[CAL16694.1] - (AL021867) putative protein [Arabidopsis thaliana]			
3128	87674330 (6255, 6256)	Novel Protein sim. GBank g13885628 (AF090133) - lin-7-A [Rattus norvegicus]	Contains protein domain (PF00595) - PDZ domain (also known as DHR or GLGF)	misc_channel	22276998, 264259, 52644045, 265008, 21906754, 265017, 265018, 21906766, 18108376, 18108387, 22279000, 22279002
3129	87755412 (6257, 6258)	Novel Protein sim. GBank g13135273 (AC003056) - hypothetical protein [Arabidopsis thaliana]	Contains protein domain (PF00400) - WD domain, G-beta repeat	kinase	58182575, 264259, 29331825, 29331826, 52644045, 58182435, 60433356, 264600, 264682, 264783, 264784, 264389, 264288, 264688, 55811957, 264682, 33657023, 33657109, 60432113, 264564, 264566, 264636
3130	14893860 (6259, 6260)	Novel Protein sim. GBank g13329485 (AF064553) - NSD1 protein [Mus musculus]			
3131	95351469 (6261, 6262)	Novel Protein sim. GBank g11648277 (U86138) - telomerase-associated protein TP-1 [Homo sapiens]	Contains protein domain (PF00400) - WD domain, G-beta repeat	UNCLASSIFIED	58182575, 264259, 29331824, 264807, 58182435, 264594, 60433438, 55812036, 33109954, 21908754, 33657084, 87168474, 264448, 264768, 21906769, 55811957, 265020, 265021, 265022, 80170615, 33657023, 33657109, 33657182, 27488261, 33657348, 85274781, 60170394, 58182323, 63373044, 87168518, 264584

3132	95415459 (6263, 6264)	Novel Protein sim. GBank g 4680847 gb AA02713.1 AF13293 - (AF13293) CG1-03 protein [Homo sapiens]	Contains protein domain (PF00789) - ubiquitin UBX domain		ubiquitin	52644507, 52646842, 52646385, 55274572, 56182575, 22276994, 22278995, 35696286, 56964075, 22278996, 22278997, 22278998, 22276999, 60432048, 52645080, 29331822, 29331824, 29331825, 80432289, 29331826, 29331827, 29331828, 35896052, 284508, 52644045, 56182435, 264810, 60170631, 60432228, 60433356, 33657402, 55812038, 52648317, 21906754, 52644296, 85656542, 87168558, 285017, 265018, 265019, 264448, 264288, 264389, 52644229, 21906765, 21906766, 21906787, 21906788, 21906769, 35695917, 265020, 265021, 60170615, 52644150, 264692, 33657023, 52845129, 33657109, 33657182, 27488261, 27488262, 27486265, 33657349, 35695783, 18108374, 18108376, 55611576, 35695855, 18108385, 18108387, 56526486, 87168518, 60432113, 22279002
3133	87379414 (6265, 6266)	Novel Protein sim. GBank g 4507613 ef NP_003738.1 pTnKS - TANKYRASE			polymerase	22278994, 22278998, 264805, 265006, 265007, 87168559, 264780, 21906767, 18108374, 22279000, 22279002, 264563 264595, 264389, 264665, 264628, 264588
3134	84648816 (6267, 6268)	Novel Protein sim. GBank g 1729827 s p P54633 TALA_DIC1 - FILIPODIN (TALIN HOMOLOG)			struct	22278996, 264085, 28331826, 33657402, 18108346, 263974
3135	86389356 (6269, 6270)	Novel Protein sim. GBank g 3053478 (AF012927) - fibrogen-binding protein (Streptococcus equi)			transport	22278998, 264259, 264628, 265006, 265008, 80433438, 265019, 284764, 264288, 264769, 264689, 265020, 27486262, 263972, 55274791, 264557, 264558
3136	84645039 (6271, 6272)	Novel Protein sim. GBank g 827101 p j S44092 - probable carrier protein c2 - Caenorhabditis elegans	Contains protein domain (PF00153) - Mitochondrial carrier proteins		UNCLASSIFIED	22278995, 22278996, 22278997, 22278998, 22278999, 264259, 28331822, 29331825, 29331826, 29331827, 29331828, 284510, 265008, 21906754, 87188474, 285011, 87168559, 265017, 265018, 265019, 18108351, 264882, 264789, 21906765, 21906766, 21906767, 21906768, 55811957, 35695917, 265020, 285021, 52644150, 18108370, 18108374, 22279000, 22279002, 264482, 264488
3137	88257847 (6273, 6274)	Novel Protein sim. GBank g 3342730 (AC005331) - R31341_1 [Homo sapiens]				

3138	94130186 (6275, 6278)	Novel Protein sim. GBank g14408759p9bAAD200701 - (AC008838) hypothetical protein [Arabidopsis thaliana]				264569, 264488, 264907, 264511, 264593, 33109854, 67188559, 264661, 264684, 264665, 264688, 264667, 264768, 264688, 264689, 264691, 264692, 264893, 33657109, 264631, 264634, 264635, 264636, 264637, 60170394, 83373044, 18106385, 18108388, 60432113, 22279000, 22279002
3139	87325503 (6277, 6278)	Novel Protein sim. GBank g1228938p11614452C - Hypo- rich glycoprotein [Zea diploperennis]			UNCLASSIFIED	22276897, 22278996, 22276899, 264905, 265018, 265019, 21908785, 265020, 264636, 264557
3140	87222892 (6278, 6280)	Novel Protein sim. GBank g19321emj[CAJ37773] - (X53744) 88kDa subunit of signal recognition particle [Carnis familiaris]			struct	22276895, 36894075, 35866286, 264908, 264908, 60433356, 21908754, 52844296, 87166474, 87188559, 264663, 284286, 264885, 264686, 265022, 264693, 27486262, 35695855, 264630, 264555, 264566
3141	87323564 (6281, 6282)	Novel Protein sim. GBank g13213227 (AF035209) - putative v-SNARE Vti1a [Mus musculus]			UNCLASSIFIED	56182575, 35696286, 28331828, 264909, 265008, 265018, 18108351, 264369, 21908768, 29148627, 265020, 264628, 264629, 264631, 18108385
3142	69418026 (6283, 6284)	Novel Protein sim. GBank g12498197[sp1Q95245[C561_P1G - CYTOCHROME B561 (CYTOCHROME B-561)]			cytochrome	52845156, 52846385, 22276895, 35696286, 22278996, 22278999, 80432049, 264259, 28331822, 29331824, 29331827, 29164499, 56182435, 265007, 80170831, 80432228, 33657402, 264595, 80433438, 284758, 21908754, 264268, 264768, 264687, 52844228, 21908785, 21906767, 21906768, 60170615, 52844150, 65274620, 33857109, 35695763, 18108370, 18108376, 65274781, 35695655, 264631, 264557, 87168518, 60432113, 22279000
3143	95351475 (6285, 6288)	Novel Protein sim. GBank g15420387[emj[CAB6679_1] - (AJ243459) proteophosphoglycan [Leishmania major]			UNCLASSIFIED	264488, 56182575, 22278996, 22278998, 22278999, 29331822, 28331824, 60432289, 35696052, 28331828, 264508, 264905, 264908, 264907, 264906, 264908, 52844045, 56182435, 264511, 264512, 265008, 264910, 60432228, 33657402, 60433356, 60433438, 55812038, 265011, 265018, 264780, 264763, 264448, 264764, 264884, 264288, 264685, 264688, 264766, 284889, 21906765, 21906766, 21906767, 21906769, 35695917, 264690, 33857023, 264683, 263967, 33657109, 264628, 264629, 18108374, 263976, 55811576, 35895855, 264630, 264631, 264632, 264634, 264635, 264636, 264837, 264558, 87168518, 60432113, 22279000, 22279002, 264563, 264566, 264486

3144	95336329 (6287, 6288)	Novel Protein sim. GBank gij4684468 emb CAB43322.1 - (AL050225) hypothetical protein [Homo sapiens]				264480, 18108398, 22278998, 35696288, 22278997, 22278999, 29331828, 29331827, 35696052, 29331828, 264106, 265006, 265007, 265009, 33857402, 85856542, 265011, 18108351, 264448, 264369, 21908765, 21908768, 21908767, 265020, 265021, 52644150, 27486281, 18108370, 18108374, 35696423, 58182323, 83373044, 22279000, 22279002, 264587
3145	88611857 (6298, 6299)	Novel Protein sim. GBank gij3879709 emb CAB03300 - (Z61118) Similarly to Human endosomal protein P182 (TR:O15075); cDNA EST EMBL:Z14467 comes from this gene; cDNA EST EMBL:Z14456 comes from this gene; cDNA EST EMBL:ID7011 comes from this gene; cDNA EST EMBL:ID7015 comes from l...			UNCLASSIFIED	18108397, 29331824, 29148499, 20281100, 265006, 55812038, 265010, 21908768, 29148627, 21908769, 29148784, 264892, 33657023, 33657109, 35695763, 263981, 58182323, 87188516
3146	87758314 (6291, 6292)	Novel Protein sim. GBank gij2135746 pf 568690 - mitogen inducible gene mlg-2 - human	Contains protein domain (PF00189) - PH domain	struct		264259, 29331828, 29331828, 29331830, 264510, 264511, 265007, 285009, 284800, 265017, 18108351, 284448, 284369, 21908768, 285021, 264892, 33857109, 18108374, 35696423, 35695855, 60432113, 264584
3147	84848512 (6293, 6294)	Novel Protein sim. GBank gij3874279 emb CAB07315.1 - (Z92825) predicted using GeneIndex; cDNA EST YK315612.3 comes from this gene; cDNA EST YK315612.5 comes from this gene [Caenorhabditis elegans]	Contains protein domain (PF00702) - halosced dehalogenase-like hydrolase	UNCLASSIFIED		56181688, 35696288, 60432048, 264259, 56182181, 29331825, 60432289, 35698052, 56182435, 285008, 264910, 60431735, 60433356, 60433438, 265010, 284448, 264288, 285022, 33857023, 33857109, 60431528, 65274791, 264631, 56182323, 264404, 22279002
3148	95362188 (6295, 6296)	Novel Protein sim. GBank gij5225322 gb AA040851.1 AF08310 - (AF083108) siruin type 3 [Homo sapiens]		UNCLASSIFIED		35698288, 35696052, 264511, 85658542, 87186474, 264704, 35696423, 284555, 264558, 264557, 264558, 83373044, 56528488, 60432113
3149	85308546 (6297, 6298)	Novel Protein sim. GBank gij4200446 (AF102777) - FYVE finger-containing phospholipidase [Mus musculus]	Contains protein domain (PF01363) - FYVE zinc finger	eph		29331822, 35696052, 264109, 29148628, 18108381
3150	87655472 (6299, 6300)	Novel Protein sim. GBank gij3378454 emb CAA78839 - (Y17850) ganglioside-induced differentiation associated protein 1 [Mus musculus]	Contains protein domain (PF00043) - Glutathione S-transferases.	transferase		284259, 29331822, 29331824, 29331825, 29331827, 52846317, 264696, 35695855, 56182323, 264639
3151	67772355 (6301, 6302)	Novel Protein sim. GBank gij172591 (M63577) - SFP1 [Saccharomyces cerevisiae]	Contains protein domain (PF00096) - Zinc finger, C2H2 type	oncogene		29331822, 265006
3152	85698108 (6303, 6304)			UNCLASSIFIED		21908754, 87186559, 264805, 21908768, 52844150, 27486284, 35696423, 22279000

3153	95317299 (6305, 6306)	Novel Protein sim. GBank g 1489504 1gb AAD32705.1 AF14395 - (AF14395) coronin-3 [Mus musculus]	Contains protein domain (PF000400) - struct	WD domain, G-beta repeat	struct	264488, 5264535, 35696286, 22278996, 22278997, 22278998, 60432049, 264259, 29331826, 60432288, 33656670, 264508, 264806, 33657402, 264595, 60433438, 87168474, 87168559, 264601, 265019, 264448, 264682, 264764, 264288, 264369, 264786, 21906765, 21906766, 21906767, 21906768, 21906769, 29146784, 265021, 265022, 60170615, 52644150, 264690, 264691, 33657023, 69274620, 33657109, 16106370, 35695855, 264638, 60170394, 87168518, 60432113, 22279000, 22279002, 22278996, 264259, 29331824, 66712502, 265008, 265010, 265017, 18108354, 264691, 33657023, 264693, 20281149, 18108374, 29331828, 264509, 264905, 264908, 264510, 264511, 264512, 33657402, 264681, 264683, 33657023, 18108370, 264634, 264639, 18108385, 264583, 264488
3154	87718573 (6307, 6308)	Novel Protein sim. GBank g 1480661 gb AAD27720.1 AF13294 - (AF13294) CGI-11 protein [Homo sapiens]			ATPase, associated	56182575, 22278996, 22278997, 22278998, 22278999, 60432049, 264259, 29331822, 29331824, 66714117, 29331825, 29331828, 29331827, 35696052, 52644045, 265007, 265009, 60170631, 60432229, 60433356, 21908754, 33108954, 67168474, 265010, 265017, 265018, 265019, 18108351, 264446, 264288, 264689, 21908768, 21908769, 21908770, 35695917, 265020, 265022, 264682, 18108370, 35696423, 56182323, 22279002
3155	87762394 (6309, 6310)	Novel Protein sim. GBank g 1728837 gb F39194 ALU7_HUMAN - IIII ALU SUBFAMILY SQ WARNING ENTRY IIII	Contains protein domain (PF00652) - transferase	Similarity to lectin domain of rdn beta-chain, 3 copies.	UNCLASSIFIED	22278996, 264259, 29331824, 66712502, 265008, 265010, 265017, 18108354, 264691, 33657023, 264693, 20281149, 18108374, 29331828, 264509, 264905, 264908, 264510, 264511, 264512, 33657402, 264681, 264683, 33657023, 18108370, 264634, 264639, 18108385, 264583, 264488
3156	87737448 (6311, 6312)	Novel Protein sim. GBank g 1563007 gb AAD45821.1 AC00601 - (AC00601) N-acetylglucosaminyltransferase: similar to Q10473 (P D g1709559) [Homo sapiens]				56182575, 22278996, 22278997, 22278998, 22278999, 60432049, 264259, 29331822, 29331824, 66714117, 29331825, 29331828, 29331827, 35696052, 52644045, 265007, 265009, 60170631, 60432229, 60433356, 21908754, 33108954, 67168474, 265010, 265017, 265018, 265019, 18108351, 264446, 264288, 264689, 21908768, 21908769, 21908770, 35695917, 265020, 265022, 264682, 18108370, 35696423, 56182323, 22279002
3157	88259577 (6313, 6314)					16108386, 264259, 29331826, 35696052, 29146498, 87168559, 265017, 264448, 264268, 264691, 18108388, 52645129, 35696423, 52644332
3158	80034118 (6315, 6316)	Novel Protein sim. GBank g 1530604 gb AAD41695.1 AF15677 - (AF15677) ASB-3 protein [Homo sapiens]	Contains protein domain (PF00023) - kinase	Ank repeat	kinase	264488, 263974
3159	84124114 (6317, 6318)	Novel Protein sim. GBank g 15531272 gb CAB50897.11 - (AJ243800) WSC4 homologue [Kluyveromyces fragilis]	UNCLASSIFIED			56182575, 22278998, 29331824, 264106, 60433356, 264758, 265011, 87168559, 264448, 18108354, 264768, 21906768, 265020, 264691, 264692, 33657109, 18108374, 35696423, 264555, 60170394, 22278996
3160	80221068 (6319, 6320)	Novel Protein sim. GBank g 13930525 (AF064447) - sex-determination protein homolog Fem1a [Mus musculus]	Contains protein domain (PF00023) - struct	Ank repeat		16108351, 264555, 264556, 264557, 264558, 264559

3161	08074111 (6321, 6322)				264408, 22278995, 22278997, 22278998, 264289, 29331822, 60432289, 29331828, 52844045, 285017, 285018, 264448, 264288, 21806764, 21806787, 285020, 18108374, 264636, 264586
------	-----------------------	--	--	--	--

Table 2

Tissue ID	Tissue Name	Tissue Information	Disease Association
20281069	192xN	Protein-protein interactions	Any
20281071	192xN	Protein-protein interactions	Any
20281149	192xN	Protein-protein interactions	Any
20281152	192xN	Protein-protein interactions	Any
264111	276xN	Protein-protein interactions	Any
264112	276xN	Protein-protein interactions	Any
263966	384xN	Protein-protein interactions	Any
263967	384xN	Protein-protein interactions	Any
264110	552xN	Protein-protein interactions	Any
18108379	5PH 52.1 (Adrenal Gland)	Adrenal Gland/Suprarenal gland	Adrenoleukodystrophy, Congenital Adrenal Hyperplasia,
18108381	5PH 52.2 (Fetal Lung)	Fetal Lung	Cystic Fibrosis, infection, lung cancer
18108383	5PH 52.3 (B's Lymphoma - Raji)	Lymphoma derived from B cells	Blood cancers, hematopoietic, leukemia
18108368	5PH 52.5 (Salivary Gland)	Salivary Gland	Dry mouth, infection
18108384	5PH 52.6 (Brain - Thalamus)	Thalamus	Brain cancer, head injury, obesity, neurological disorders,
18108394	5PH 53.1 (Adrenal Gland)	Adrenal Gland/Suprarenal gland	Adrenoleukodystrophy, Congenital Adrenal Hyperplasia,
18108355	5PH 53.2 (Fetal Lung)	Fetal Lung	Cystic Fibrosis, infection, lung cancer
18108359	5PH 53.3 (B's Lymphoma - Raji)	Lymphoma derived from B cells	Blood cancers, hematopoietic, leukemia
18108361	5PH 53.4 (Mammary Gland)	Mammary Gland	Lactation disorders, breast cancer
18108362	5PH 53.5 (Salivary Gland)	Salivary Gland	Dry mouth, infection
18108366	5PH 53.6 (Brain - Thalamus)	Thalamus	Brain cancer, head injury, obesity, neurological disorders,
18108354	5PH 54.1 (Adrenal Gland)	Adrenal Gland/Suprarenal gland	Adrenoleukodystrophy, Congenital Adrenal Hyperplasia,
18108392	5PH 54.2 (Fetal Lung)	Fetal Lung	Cystic Fibrosis, infection, lung cancer
18108348	5PH 54.3 (B's Lymphoma - Raji)	Lymphoma derived from B cells	Blood cancers, hematopoietic, leukemia
18108382	5PH 54.4 (Mammary Gland)	Mammary Gland	Lactation disorders, breast cancer
18108395	5PH 54.5 (Salivary Gland)	Salivary Gland	Dry mouth, infection
18108365	5PH 54.6 (Brain - Thalamus)	Thalamus	Brain cancer, head injury, obesity, neurological disorders,
18108397	5PH 55.1 (Adrenal Gland)	Adrenal Gland/Suprarenal gland	Adrenoleukodystrophy, Congenital Adrenal Hyperplasia,
18108398	5PH 55.2 (Fetal Lung)	Fetal Lung	Cystic Fibrosis, infection, lung cancer
18108364	5PH 55.3 (B's Lymphoma - Raji)	Lymphoma derived from B cells	Blood cancers, hematopoietic, leukemia
18108388	5PH 55.4 (Mammary Gland)	Mammary Gland	Lactation disorders, breast cancer
18108358	5PH 55.5 (Salivary Gland)	Salivary Gland	Dry mouth, infection
20281099	5PH 56.2 (MIG63)		
20281100	5PH 56.3 (UISMC)		
264404	5PH 1 (Brain)	Whole Brain	Von Hippel-Lindau (VHL) syndrome, Alzheimer's disease, Stroke, Tuberculous sclerosis, hypercalcaemia, Parkinson's disease, Huntington's disease, Cerebral palsy, Epilepsy, Lesch-Nyhan syndrome, Multiple sclerosis, Ataxia-telangiectasia, Leukodystrophies, Behavioral disorders, Addiction, Anxiety, Pain, Neuroprotection

264510	SPH.10 (Pancreas)	Pancreas	Pancreatitis, diabetes, pancreatic cancer
264511	SPH.11 (Placenta)	Placenta	Infertility, birth defects
264512	SPH.12 (Thyroid)	Thyroid	Hypoparathyroidism, Hypoparathyroidism
264555	SPH.13 (Bone Marrow)	Bone Marrow	Hemophilia, hypercoagulation, idiopathic thrombocytopenic purpura, autoimmune disease, allergies, immunodeficiencies, transplantation, Graft versus host.
264556	SPH.14 (Bone Marrow)	Bone Marrow	Hemophilia, hypercoagulation, idiopathic thrombocytopenic purpura, autoimmune disease, allergies, immunodeficiencies, transplantation, Graft versus host.
264557	SPH.15 (Bone Marrow)	Bone Marrow	Hemophilia, hypercoagulation, idiopathic thrombocytopenic purpura, autoimmune disease, allergies, immunodeficiencies, transplantation, Graft versus host.
264558	SPH.16 (Bone Marrow)	Bone Marrow	Hemophilia, hypercoagulation, idiopathic thrombocytopenic purpura, autoimmune disease, allergies, immunodeficiencies, transplantation, Graft versus host.
264559	SPH.17 (Bone Marrow)	Bone Marrow	Hemophilia, hypercoagulation, idiopathic thrombocytopenic purpura, autoimmune disease, allergies, immunodeficiencies, transplantation, Graft versus host.
264569	SPH.19 (One Fetal tissue and two cell lines)	Mixed	
264687	SPH.19.1 (Fetal thymus - CRL 7046)	Fetal Thymus	Hemophilia, hypercoagulation, idiopathic thrombocytopenic purpura, immunodeficiencies
264688	SPH.19.2 (Hematopoietic stem cells - CRL 2043)	Hematopoietic stem cells	Leukemia, osteoporosis, post-chemotherapy stem cell repopulation
264689	SPH.19.3 (osteogenic sarcoma cell lines - HTB96)	Osteogenic Sarcoma	Sarcomas, osteoporosis, osteopetrosis
264690	SPH.19.4 (Fetal Liver)	Fetal liver	Von Hippel-Lindau (VHL) syndrome, Cirrhosis, Transplantation
264691	SPH.19.5 (Heart)	Heart	Cardiomyopathy, Atherosclerosis, Hypertension, Congenital heart defects, Aortic stenosis, Atrial septal defect (ASD), Atrioventricular (A-V) canal defect, Ductus arteriosus, Pulmonary stenosis, Subaortic stenosis, Ventricular septal defect (VSD), valve diseases, Tuberosclerosis, Scleroderma, Obesity, Transplantation
264692	SPH.19.6 (Spleen)	Spleen	Hemophilia, Hypercoagulation, Idiopathic thrombocytopenic purpura, Immunodeficiencies, Graft versus host
264693	SPH.19.7 (Pituitary)	Pituitary	Von Hippel-Lindau (VHL) syndrome, Alzheimer's disease, Stroke, Tuberosclerosis, hypercalcaemia, Parkinson's disease, Huntington's disease, Cerebral palsy, Epilepsy, Lesch-Nyhan syndrome, Multiple sclerosis, Ataxia-telangiectasia, Leukodystrophies, Behavioral disorders, Addition, Anxiety, Pain, Neuropathic
264482	SPH.2 (Brain)	Brain	Von Hippel-Lindau (VHL) syndrome, Alzheimer's disease, Stroke, Tuberosclerosis, hypercalcaemia, Parkinson's disease, Huntington's disease, Cerebral palsy, Epilepsy, Lesch-Nyhan syndrome, Multiple sclerosis, Ataxia-telangiectasia, Leukodystrophies, Behavioral disorders, Addition, Anxiety, Pain, Neuropathic

264600	SPH.21 (Fetal Brain)	Fetal brain	Von Hippel-Lindau (VHL) syndrome, Alzheimer's disease, Stroke, Tuberculous sclerosis, hypercalcaemia, Parkinson's disease, Huntington's disease, Cerebral palsy, Epilepsy, Lesch-Nyhan syndrome, Multiple sclerosis, Ataxia-telangiectasia, Leukodystrophies, Behavioral disorders, Addiction, Anxiety, Pain, Neuroprotection
264601	SPH.22 (Bone Marrow)	Bone Marrow	Hemophilia, hypercoagulation, idiopathic thrombocytopenic purpura, autoimmune disease, allergies, immunodeficiencies, transplantation, Graft versus host.
264602	SPH.23 (Thyroid)	Thyroid	Hyperthyroidism and Hypothyroidism
264603	SPH.24 (Pancreas)	Pancreas	Pancreatitis, diabetes, pancreatic cancer
264604	SPH.25 (Lymph Node)	Lymph Node	Lymphedema, Allergies
264605	SPH.26 (Placenta)	Placenta	Infertility, birth defects
264634	SPH.28 (Heart)	Heart	Cardiomyopathy, Atherosclerosis, Hypertension, Congenital heart defects, Aortic stenosis, Atrial septal defect (ASD), Atrioventricular (A-V) canal defect, Ductus arteriosus, Pulmonary stenosis, Subaortic stenosis, Ventricular septal defect (VSD), valve diseases, Tuberculous sclerosis, Scleroderma, Obesity, Transplantation
264635	SPH.29 (Fetal Kidney)	Fetal Kidney	Diabetes, Autoimmune disease, Renal artery stenosis, Interstitial nephritis, Glomerulonephritis, Polycystic kidney disease, Systemic lupus erythematosus, Renal tubular acidosis, IgA nephropathy, Hypercalcaemia, Lesch-Nyhan syndrome
264636	SPH.3 (Bone Marrow)	Bone marrow	Hemophilia, hypercoagulation, idiopathic thrombocytopenic purpura, autoimmune disease, allergies, immunodeficiencies, transplantation, Graft versus host.
264637	SPH.30 (Lymph Node)	Lymph Node	Lymphedema, Allergies
264638	SPH.31 (Pancreas)	Pancreas	Pancreatitis, diabetes, pancreatic cancer
264639	SPH.32 (Thyroid)	Thyroid	Hyperthyroidism and Hypothyroidism
264639	SPH.33 (Fetal Brain)	Fetal brain	Von Hippel-Lindau (VHL) syndrome, Alzheimer's disease, Stroke, Tuberculous sclerosis, hypercalcaemia, Parkinson's disease, Huntington's disease, Cerebral palsy, Epilepsy, Lesch-Nyhan syndrome, Multiple sclerosis, Ataxia-telangiectasia, Leukodystrophies, Behavioral disorders, Addiction, Anxiety, Pain, Neuroprotection
264484	SPH.4 (Bone Marrow)	Bone Marrow	Hemophilia, hypercoagulation, idiopathic thrombocytopenic purpura, autoimmune disease, allergies, immunodeficiencies, transplantation, Graft versus host.
264758	SPH.44.1 (Kidney)	Kidney	Diabetes, Autoimmune disease, Renal artery stenosis, Interstitial nephritis, Glomerulonephritis, Polycystic kidney disease, Systemic lupus erythematosus, Renal tubular acidosis, IgA nephropathy, Hypercalcaemia, Lesch-Nyhan syndrome
264760	SPH.44.2 (Fetal Liver)	Fetal Liver	Von Hippel-Lindau (VHL) syndrome, Cirrhosis, Transplantation

264762	SPH.44.3 (Heart)	Heart	Cardiomyopathy, Atherosclerosis, Hypertension, Congenital heart defects, Aortic stenosis, Atrial septal defect (ASD), Atrioventricular (A-V) canal defect, Ductus arteriosus, Pulmonary stenosis, Subaortic stenosis, Ventricular septal defect (VSD), valve diseases, Tuberosclerosis, Scleroderma, Obesity, Transplantation
264764	SPH.44.4 (Prostate)	Prostate Cancer	
264766	SPH.44.5 (Spleen)	Spleen	Hemophilia, Hypocoagulation, Idiopathic thrombocytopenic purpura, Immunodeficiencies, Graft versus host
264768	SPH.44.6 (pituitary)	Pituitary	Von Hippel-Lindau (VHL) syndrome, Alzheimer's disease, Stroke, Tuberosclerosis, hypercalcaemia, Parkinson's disease, Huntington's disease, Cerebral palsy, Epilepsy, Lesch-Nyhan syndrome, Multiple sclerosis, Ataxia-sclerostasis, Leukodystrophies, Behavioral disorders, Addiction, Anxiety, Pain, Neuroprotection
264769	SPH.44.7 (Uterus)	Uterus	Infertility, birth defects
264905	SPH.48.1 (Burkitt's Lymphoma-Raji)	Burkitt's Lymphoma	Lymphoma, blood cancers
264906	SPH.48.2 (Thalamus-Brain)	Thalamus	Von Hippel-Lindau (VHL) syndrome, Alzheimer's disease, Stroke, Tuberosclerosis, hypercalcaemia, Parkinson's disease, Huntington's disease, Cerebral palsy, Epilepsy, Lesch-Nyhan syndrome, Multiple sclerosis, Ataxia-sclerostasis, Leukodystrophies, Behavioral disorders, Addiction, Anxiety, Pain, Neuroprotection
264907	SPH.48.3 (Adrenal Gland)	Adrenal Gland/Suprarenal gland	Adrenoleukodystrophy, Congenital Adrenal Hyperplasia,
264908	SPH.48.4 (Fetal Lung)	Fetal Lung	Cystic Fibrosis, infection, lung cancer
264909	SPH.48.5 (Salivary Gland)	Salivary Gland	Dry mouth, infection
264910	SPH.48.6 (Mammary Gland)	Mammary Gland	Lactation disorders, breast cancer
265006	SPH.50.1 (B's Lymphoma)	Burkitt's Lymphoma	Lymphoma, blood cancers
265007	SPH.50.2 (thalamus)	Thalamus	Von Hippel-Lindau (VHL) syndrome, Alzheimer's disease, Stroke, Tuberosclerosis, hypercalcaemia, Parkinson's disease, Huntington's disease, Cerebral palsy, Epilepsy, Lesch-Nyhan syndrome, Multiple sclerosis, Ataxia-sclerostasis, Leukodystrophies, Behavioral disorders, Addiction, Anxiety, Pain, Neuroprotection
265008	SPH.50.3 (adrenal gland)	Adrenal Gland/Suprarenal gland	Adrenoleukodystrophy, Congenital Adrenal Hyperplasia,
265009	SPH.50.4 (fetal lung)	Fetal Lung	Cystic Fibrosis, infection, lung cancer
265010	SPH.50.5 (salivary gland)	Salivary Gland	Dry mouth, infection
265011	SPH.50.6 (mammary gland)	Mammary Gland	Lactation disorders, breast cancer
18108385	SPH.51.1 (MCF-7)	Breast Cancer	Breast Cancer
18108370	SPH.51.2 (CCRF-CEM)	Cancer Cell line	Cancer
18108374	SPH.51.3 (K-562)	Cancer Cell line	Cancer
18108351	SPH.51.4 (OVCAR-3)	Ovarian cancer	Cancer
18108372	SPH.51.5 (HL-60)	Cancer Cell line	Cancer
264486	SPH.6 (Bone Marrow)	Bone Marrow	Hemophilia, hypocoagulation, idiopathic thrombocytopenic purpura, autoimmune diseases, allergies, immunodeficiencies, transplantation, Graft versus host,

264508	5PH.8 (Fetal Brain)	Von Hippel-Lindau (VHL) syndrome, Alzheimer's disease, Stroke, Tuberculous sclerosis, hypercalcaemia, Parkinson's disease, Huntington's disease, Cerebral palsy, Epilepsy, Lesch-Nyhan syndrome, Multiple sclerosis, Ataxia-teleangiectasia, Leukodystrophies, Behavioral disorders, Addiction, Anxiety, Pain, Neuroprotection	
264509	5PH.9 (Lymph Node)	Lymph Node	Lymphedema, Allergies
20798451	5RH.56.3 (UIMC)		
264487	5RH.1 (Brain)	Brain	Von Hippel-Lindau (VHL) syndrome, Alzheimer's disease, Stroke, Tuberculous sclerosis, hypercalcaemia, Parkinson's disease, Huntington's disease, Cerebral palsy, Epilepsy, Lesch-Nyhan syndrome, Multiple sclerosis, Ataxia-teleangiectasia, Leukodystrophies, Behavioral disorders, Addiction, Anxiety, Pain, Neuroprotection
264534	5RH.11 (Bone marrow)	Bone Marrow	Hemophilia, hypercoagulation, idiopathic thrombocytopenic purpura, autoimmune disease, allergies, immunodeficiencies, transplantation, Graft versus host,
264535	5RH.12 (Bone marrow)	Bone Marrow	Hemophilia, hypercoagulation, idiopathic thrombocytopenic purpura, autoimmune disease, allergies, immunodeficiencies, transplantation, Graft versus host,
264563	5RH.19 (Fetal Brain)	Fetal brain	Von Hippel-Lindau (VHL) syndrome, Alzheimer's disease, Stroke, Tuberculous sclerosis, hypercalcaemia, Parkinson's disease, Huntington's disease, Cerebral palsy, Epilepsy, Lesch-Nyhan syndrome, Multiple sclerosis, Ataxia-teleangiectasia, Leukodystrophies, Behavioral disorders, Addiction, Anxiety, Pain, Neuroprotection
264488	5RH.2 (Bone Marrow)	Bone Marrow	Hemophilia, hypercoagulation, idiopathic thrombocytopenic purpura, autoimmune disease, allergies, immunodeficiencies, transplantation, Graft versus host,
264564	5RH.20 (Lymph Node)	Lymph Node	Lymphedema, Allergies
264565	5RH.21 (Pancreas)	Pancreas	Pancreatitis, diabetes, pancreatic cancer
264566	5RH.22 (Placenta)	Placenta	Infertility, birth defects
264567	5RH.23 (Thyroid)	Thyroid	Hyperthyroidism and Hypothyroidism
264591	5RH.25 (Fetal Brain)	Fetal brain	Von Hippel-Lindau (VHL) syndrome, Alzheimer's disease, Stroke, Tuberculous sclerosis, hypercalcaemia, Parkinson's disease, Huntington's disease, Cerebral palsy, Epilepsy, Lesch-Nyhan syndrome, Multiple sclerosis, Ataxia-teleangiectasia, Leukodystrophies, Behavioral disorders, Addiction, Anxiety, Pain, Neuroprotection
264592	5RH.26 (Bone Marrow)	Bone Marrow	Hemophilia, hypercoagulation, idiopathic thrombocytopenic purpura, autoimmune disease, allergies, immunodeficiencies, transplantation, Graft versus host,
264593	5RH.27 (Thyroid)	Thyroid	Hyperthyroidism and Hypothyroidism
264594	5RH.28 (Pancreas)	Pancreas	Pancreatitis, diabetes, pancreatic cancer
264595	5RH.29 (Lymph Node)	Lymph Node	Lymphedema, Allergies
264489	5RH.3 (Bone Marrow)	Bone Marrow	Hemophilia, hypercoagulation, idiopathic thrombocytopenic purpura, autoimmune disease, allergies, immunodeficiencies, transplantation, Graft versus host,

264596	SRH.30 (Placenta)	Placenta	Infertility, birth defects
264628	SRH.33 (fetal Kidney)	Fetal kidney	Diabetes, Autoimmune disease, Renal artery stenosis, Interstitial nephritis, Glomerulonephritis, Polycystic kidney disease, Systemic lupus erythematosus, Renal tubular acidosis, IgA nephropathy, Hypercalcaemia, Lesch-Nyhan syndrome
264629	SRH.34 (Lymph Node)	Lymph Node	Lymphedema, Allergies
264630	SRH.35 (Pancreas)	Pancreas	Pancreatitis, diabetes, pancreatic cancer
264631	SRH.36 (thyroid)	Thyroid	Hypothyroidism and Hypothyroidism
264632	SRH.37 (Fetal Brain)	Fetal Brain	Von Hippel-Lindau (VHL) syndrome, Alzheimer's disease, Stroke, Tuberculous sclerosis, hypercalcaemia, Parkinson's disease, Huntington's disease, Cerebral palsy, Epilepsy, Lesch-Nyhan syndrome, Multiple sclerosis, Ataxia-telangiectasia, Leukodystrophies, Behavioral disorders, Addiction, Anxiety, Pain, Neuroprotection
264490	SRH.4 (Bone Marrow)	Bone Marrow	Hemophilia, hypercoagulation, idiopathic thrombocytopenic purpura, autoimmune disease, allergies, immunodeficiencies, transplantation, Graft versus host.
264681	SRH.43.1 (fetal thymus - CRL7046)	Fetal Thymus	Hemophilia, hypercoagulation, idiopathic thrombocytopenic purpura, immunodeficiencies
264682	SRH.43.2 (hematopoietic stem cells - CRL2043)	Hematopoietic stem cells	Leukemia, osteoporosis, post-chemotherapy stem cell repopulation
264683	SRH.43.3 (osteogenic sarcoma cell lines - HTB96)	Osteogenic Sarcoma	Sarcomas, osteoporosis, osteopetrosis
264684	SRH.43.4 (Fetal Liver)	Fetal Liver	Von Hippel-Lindau (VHL) syndrome, Cirrhosis, Transplantation
264685	SRH.43.6 (Spleen)	Spleen	Hemophilia, Hypercoagulation, idiopathic thrombocytopenic purpura, immunodeficiencies, Graft versus host
264686	SRH.43.7 (pituitary)	Pituitary	Von Hippel-Lindau (VHL) syndrome, Alzheimer's disease, Stroke, Tuberculous sclerosis, hypercalcaemia, Parkinson's disease, Huntington's disease, Cerebral palsy, Epilepsy, Lesch-Nyhan syndrome, Multiple sclerosis, Ataxia-telangiectasia, Leukodystrophies, Behavioral disorders, Addiction, Anxiety, Pain, Neuroprotection
264757	SRH.44.1 (Kidney)	Kidney	Diabetes, Autoimmune disease, Renal artery stenosis, Interstitial nephritis, Glomerulonephritis, Polycystic kidney disease, Systemic lupus erythematosus, Renal tubular acidosis, IgA nephropathy, Hypercalcaemia, Lesch-Nyhan syndrome
264759	SRH.44.2 (Fetal Liver)	Fetal Liver	Von Hippel-Lindau (VHL) syndrome, Cirrhosis, Transplantation
264761	SRH.44.3 (Heart)	Heart	Cardiomyopathy, Atherosclerosis, Hypertension, Congenital heart defects, Aortic stenosis, Atrial septal defect (ASD), Atrioventricular (A-V) canal defect, Ductus arteriosus, Pulmonary stenosis, Subaortic stenosis, Ventricular septal defect (VSD), valve diseases, Tuberculous sclerosis, Scleroderma, Obesity, Transplantation
264763	SRH.44.4 (Prostate)	Prostate	Prostate Cancer
264765	SRH.44.5 (Spleen)	Spleen	Hemophilia, Hypercoagulation, idiopathic thrombocytopenic purpura, immunodeficiencies, Graft versus host

264767	SRH.44.6 (Pituitary)	Pituitary	Von Hippel-Lindau (VHL) syndrome, Alzheimer's disease, Stroke, Tuberculous sclerosis, hypercalcaemia, Parkinson's disease, Huntington's disease, Cerebral palsy, Epilepsy, Lesch-Nyhan syndrome, Multiple sclerosis, Ataxia-telangiectasia, Leukodystrophies, Behavioral disorders, Addiction, Anxiety, Pain, Neuroprotection
264828	SRH.46.1 (Lymph Node)	Lymph Node	Lymphedema, Allergies
264887	SRH.47.5 (Fetal Liver)	Fetal Liver	Von Hippel-Lindau (VHL) syndrome, Cirrhosis, Transplantation
18108377	SRH.50.1 (B's lymphoma)	Burkitt's lymphoma	Lymphoma, blood cancers
18108380	SRH.50.2 (thalamus)	Thalamus	Von Hippel-Lindau (VHL) syndrome, Alzheimer's disease, Stroke, Tuberculous sclerosis, hypercalcaemia, Parkinson's disease, Huntington's disease, Cerebral palsy, Epilepsy, Lesch-Nyhan syndrome, Multiple sclerosis, Ataxia-telangiectasia, Leukodystrophies, Behavioral disorders, Addiction, Anxiety, Pain, Neuroprotection
18108396	SRH.50.3 (adrenal gland)	Adrenal Gland/Suprarenal gland	Adrenoleukodystrophy, Congenital Hyperplasia
18108397	SRH.50.4 (fetal lung)	Fetal Lung	Airway diseases, infection
18108357	SRH.50.5 (salivary gland)	Salivary Gland	Dry mouth, infection
18108390	SRH.50.6 (mammary gland)	Mammary Gland	Lactation disorders, breast cancer
264532	SRH.9 (Bone Marrow)	Bone Marrow	Hemophilia, hypercoagulable, idiopathic thrombocytopenic purpura, autoimmune diseases, allergies, immunodeficiencies, transplantation, Graft versus host,
263974	736xN		
263976	736xN		
263981	736xN		
20281166	96xN		
20281169	96xN		
20281171	96xN		
263994	CDNA-ORF Selection		
264080	MX96		
21906754	NQH 6.1 (HH729)		
22278996	NQH 6.10 (PPEC)	Endothelial cells	heart disease, cancer
22278997	NQH 6.11 (CAEC)	Endothelial cells	heart disease, cancer
22278998	NQH 6.12 (CSC)	Cancer Cell line	Cancer
22278999	NQH 6.13 (NHNPFC)	Cancer Cell line	Cancer
22279000	NQH 6.14 (NHMC-RM)	Cancer Cell line	Cancer
22279002	NQH 6.15 (Hypothalamus)	Hypothalamus	Von Hippel-Lindau (VHL) syndrome, Alzheimer's disease, Stroke, Tuberculous sclerosis, hypercalcaemia, Parkinson's disease, Huntington's disease, Cerebral palsy, Epilepsy, Lesch-Nyhan syndrome, Multiple sclerosis, Ataxia-telangiectasia, Leukodystrophies, Behavioral disorders, Addiction, Anxiety, Pain, Neuroprotection, Obesity
21906764	NQH 6.2 (In Dated Platelets)	Platelets	Clotting diseases, stroke
21906765	NQH 6.3 (HuVec)	Endothelial cells	heart disease, cancer
87168474	NQH 6.3 (Sized-HUVEC)	Endothelial cells	heart disease, cancer
21906766	NQH 6.4 (UMVEC-myo)	Cancer Cell line	Cancer
21906767	NQH 6.5 (NHMC-neo)	Cancer Cell line	Cancer
21906768	NQH 6.6 (NHK)	Cancer Cell line	Cancer
21906769	NQH 6.7 (BYCAEC)	Endothelial cells	heart disease, cancer
22278994	NQH 6.8 (NHA)	Cancer Cell line	Cancer

22278995	NQH 6.9 (PISC)	Cancer Cell line	Cancer
27486261	NQH 7.1 (Jurkat E6-unreated)	Cancer Cell line	Cancer
27486262	NQH 7.2 (TF1-unreated)	Cancer Cell line	Cancer
27486264	NQH 7.3 (U87-unreated)	Cancer Cell line	Cancer
27486265	NQH 7.4 (THP1-unreated)	Cancer Cell line	Cancer
29331822	NQH 8.1 (Brain-amygdala)		
29331824	NQH 8.2 (Brain-hippocampus)		
29331825	NQH 8.3 (Brain-substantia nigra)		
29331826	NQH 8.4 (small intestine)	Small intestine	digestive diseases, obesity, diabetes Addiction, Anxiety, Pain, Neuroprotection Von Hippel-Lindau (VHL) syndrome, Alzheimer's disease, Stroke, Tuberosus sclerosis, hypercalcaemia, Parkinson's disease, Huntington's disease, Cerebral palsy, Epilepsy, Lesch-Nyhan syndrome, Multiple sclerosis, Ataxia- relangictasias, Leukodystrophies, Behavioral disorders, Addition, Anxiety, Pain, Neuroprotection
29331827	NQH 8.5 (Spinal cord)	Spinal chord	digestive diseases, obesity, diabetes Addiction, Anxiety, Pain, Neuroprotection
29331828	NQH 8.6 (Stomach)	Stomach	Stomach cancer
29331830	NQH 8.7 (Trachea)	Trachea	Airway diseases, infection
87168518	NQH 9.1 (Sized-MG-63 treatment pool)		
87168559	NQH 9.2 (Sized-HEPG2-unreated)		
35695763	NQH 10.1 (MCF-7-unreated)	Cancer Cell line	Cancer
35695855	NQH 10.2 (U-937_treatment pool)	Cancer Cell line	Cancer
35695917	NQH 10.3 (JAR)	Cancer Cell line	Cancer
35696052	NQH 10.4 (PA-1)	Cancer Cell line	Cancer
35696286	NQH 10.5 (CADMEC)	Endothelial cells	heart disease, cancer
35696423	NQH 10.6 (CADMEC_LA)	Endothelial cells	heart disease, cancer
52644045	NQH 11.1 (SK-PN-DW)	Cancer Cell line	Cancer
52644150	NQH 11.2 (Chononic Villus Cells)	Chononic villus	fertility, birth defects
52644229	NQH 11.3 (A549)	Cancer Cell line	Cancer
52644296	NQH 11.4 (U266B1)	Cancer Cell line	Cancer
52644332	NQH 11.5 (Baoy)	Cancer Cell line	Cancer
52644507	NQH 11.6 (SW1783)	Cancer Cell line	Cancer
52645080	NQH 12.1 (U-118MG)	Cancer Cell line	Cancer
52645129	NQH 12.2 (A204)	Cancer Cell line	Cancer
52645156	NQH 12.3 (T24)	Cancer Cell line	Cancer
52646317	NQH 12.4 (G-401)	Cancer Cell line	Cancer
52646365	NQH 12.5 (Caski)	Cancer Cell line	Cancer
52646842	NQH 12.6 (SHP-77)	Cancer Cell line	Cancer

60424179	NQH.14.1 (Yale75_breast carcinoma)	Breast carcinoma	Breast Cancer
60424269	NQH.14.2 (Yale78B_ovar.tumor)	Ovary tumor	Ovarian cancer
60431528	NQH.14.3 (Yale79_prostateBPH)	Prostate	Prostate Cancer
60431602	NQH.14.4 (Yale79_prostateBPH)	Prostate	Prostate Cancer
60431735	NQH.14.5 (Yale80_ProstateAdenocarcinoma)	Prostate	Prostate Cancer
60431850	NQH.14.6 (Yale86_UterineMyoma)	Uterine Myoma	Uterine Cancer
60432049	NQH.15.1 (Yale99_cervix)	Cervix	Fertility
60432113	NQH.15.2 (Yale45_spleenITP)	Cervix	Osteoporosis, cervical cancer
60432229	NQH.15.3 (Yale16_Skin)	Skin	Hemophilia, Hypertension, Idiopathic thrombocytopenic purpura, Immunodeficiencies, Graft versus host
60432289	NQH.15.4 (Yale137_Parotid)	Skin	wound healing, melanoma
60433356	NQH.15.5 (Yale38_SmallIntestine)	Small intestine	digestive diseases, obesity, diabetes
60433438	NQH.15.6 (Yale28_ColonAscending)	Colon	Colon cancer
65274444	NQH.17.1 (Larynx)	Larynx	Cancer
65274572	NQH.17.2 (Duodenum)	Duodenum	Cancer
65274620	NQH.17.3 (Kidney, Primary tumors)	Kidney	Diabetes, Autoimmune disease, Renal artery stenosis, Interstitial nephritis, Glomerulonephritis, Polycystic kidney disease, Systemic lupus erythematosus, Renal tubular acidosis, IgA nephropathy, Hypertension, Lesch-Nyhan syndrome
65274727	NQH.17.4 (Lung Pleura, normal)	Lung	Airway diseases, infection
65274791	NQH.17.5 (Lung, Normal Adult)	Lung	Airway diseases, infection
83373044	NQH.18.230 (Pooled adrenal gland, placenta)	Adrenal Gland/Suprarenal gland	Adrenoleukodystrophy, Congenital Adrenal Hyperplasia,
85658542	NQH.18.560 (Pooled uterus, BeWo pool)	Uterus	Infertility, birth defects
33656970	NQH.9.1 (MG-63_treatment pool)	Cancer Cell line	Cancer
33657023	NQH.9.2 (HEPG2 untreated)	Cancer Cell line	Von Hippel-Lindau (VHL) syndrome, Cirrhosis, Transplantation
33657084	NQH.9.3 (PC3 untreated)	Cancer Cell line	Cancer
33657109	NQH.9.4 (TF-1_TPA)	Cancer Cell line	Cancer
33657182	NQH.9.5 (TF-1_TPO)	Cancer Cell line	Cancer
33657349	NQH.9.6 (TF-1_Hemim)	Cancer Cell line	Cancer
33657402	NQH.9.7 (HFDPC)	Cancer Cell line	Cancer
264259	NQH.1 (Mixture of eight adult & two fetal tissues)	Cancer Cell line	Cancer
264288	NQH.2 (Ten tissues plus lymphocyte control)	Cancer Cell line	Cancer
264448	NQH.3 (Bone Marrow)	Bone Marrow	Hemophilia, hypercoagulation, idiopathic thrombocytopenic purpura, autoimmune disease, allergies, immunodeficiencies, transplantation, Graft versus host,
265017	NQH.4.1 (lymph node)	Lymph Node	Lymphedema, Allergies

265018	NQ4.2 (fetal kidney)	Fetal Kidney	Diabetes, Autoimmune disease, Renal artery stenosis, Interstitial nephritis, Glomerulonephritis, Polycystic kidney disease, Systemic lupus erythematosus, Renal tubular acidosis, IgA nephropathy, Hypercalcaemia, Leish-Nyhan syndrome
66712502	NQ4.2 (Sized)		
265019	NQ4.3 (pituitary gland)		Von Hippel-Lindau (VHL) syndrome, Alzheimer's disease, Stroke, Tuberculous sclerosis, hypercalcaemia, Parkinson's disease, Huntington's disease, Cerebral palsy, Epilepsy, Leish-Nyhan syndrome, Multiple sclerosis, A-taxia-relangictasia, Leukodystrophies, Behavioral disorders, Addiction, Anxiety, Pain, Neuropathic, Obesity
66714117	NQ4.3 (Sized)		
265020	NQ4.4 (testis)	testis	Infertility, birth defects
265021	NQ4.5 (fetal liver)	Fetal Liver	Von Hippel-Lindau (VHL) syndrome, Cirrhosis, Transplantation
265022	NQ4.6 (thyroid)	Thyroid	Hypothyroidism and Hypothyroidism
18108376	NQ5.1 (MCF-7)	Breast cancer	Breast Cancer
18108387	NQ5.2 (CCRF-CEM)	Cancer Cell line	Cancer
264952	NRL1: HPLC		
	FRACTIONATION OF RE-LIG		
263971	Old BB3 Bais		
263969	Old BB5 Bais		
263975	ORFSET		
263972	OTHER Bais		
263978	PGALORF		
264106	PBBATS		
264088	QC-YA7		
264089	QC-YA8		
264102	Resequenced Interactors		
264369	RMH.1		
60170394	RMH.10.1 (MCF-7untreated)	Breast cancer	Breast Cancer
60170615	RMH.10.2 (U-937_treatment pool)	Cancer Cell line	Cancer
60170831	RMH.10.3 (JAR)	Cancer Cell line	Cancer
60174639	RMH.11.8 (Hela)	Cancer Cell line	Cancer
264113	mqEA Bais		
263973	RQEA_B5 bais		
29146498	SRD 3.1 (SKMC)	Cancer Cell line	Cancer
29146499	SRD 3.2 (SKMC)	Cancer Cell line	Cancer
29147620	SRD 3.3 (RFTC)	Cancer Cell line	Cancer
29148627	SRD 3.4 (HRCE)	Cancer Cell line	Cancer
29148629	SRD 3.6 (HRE)	Cancer Cell line	Cancer
29148784	SRD 3.7 (HRE)	Cancer Cell line	Cancer
55810764	SRD.7.1 (Lymph Node)	Lymph Node	Lymphedema, Allergies
55811150	SRD.7.2 (pancreas)	Pancreas	Pancreatitis, diabetes, pancreatic cancer
55811386	SRD.7.3 (Adrenal Gland)	Adrenal Gland/Suprarenal gland	Adrenoleukodystrophy, Congenital Adrenal Hyperplasia,

55811576	SRD.7.4 (Pituitary Gland)	Pituitary	Von Hippel-Lindau (VHL) syndrome, Alzheimer's disease, Stroke, Tuberculous sclerosis, hypercalcaemia, Parkinson's disease, Huntington's disease, Cerebral palsy, Epilepsy, Lesch-Nyhan syndrome, Multiple sclerosis, Ataxia-telangiectasia, Leukodystrophies, Behavioral disorders, Addiction, Anxiety, Pain, Neuropathology, Obesity	
55811957	SRD.7.5 (Fetal Liver)	Fetal Liver	Von Hippel-Lindau (VHL) syndrome, Cirrhosis, Transplantation	
55812038	SRD.7.6 (Fetal Kidney)	Fetal kidney	Diabetes, Autoimmune disease, Renal artery stenosis, Interstitial nephritis, Glomerulonephritis, Polycystic kidney disease, Systemic lupus erythematosus, Renal tubular acidosis, IgA nephropathy, Hypercalcaemia, Lesch-Nyhan syndrome	
56181562	SRD.8.1 (Lymph Node)	Lymph Node	Lymphedema, Allergies	
56181686	SRD.8.2 (Pancreas)	Pancreas	Pancreatitis, diabetes, pancreatic cancer	
56182181	SRD.8.3 (Adrenal Gland)	Adrenal Gland/Suprarenal gland	Adrenoleukodystrophy, Congenital Adrenal Hyperplasia,	
56182323	SRD.8.4 (Pituitary Gland)	Pituitary	Von Hippel-Lindau (VHL) syndrome, Alzheimer's disease, Stroke, Tuberculous sclerosis, hypercalcaemia, Parkinson's disease, Huntington's disease, Cerebral palsy, Epilepsy, Lesch-Nyhan syndrome, Multiple sclerosis, Ataxia-telangiectasia, Leukodystrophies, Behavioral disorders, Addiction, Anxiety, Pain, Neuropathology, Obesity	
56182435	SRD.8.5 (Fetal Liver)	Fetal Liver	Von Hippel-Lindau (VHL) syndrome, Cirrhosis, Transplantation	
56182575	SRD.8.6 (Fetal Kidney)		Diabetes, Autoimmune disease, Renal artery stenosis, Interstitial nephritis, Glomerulonephritis, Polycystic kidney disease, Systemic lupus erythematosus, Renal tubular acidosis, IgA nephropathy, Hypercalcaemia, Lesch-Nyhan syndrome	
32833986	SRD4: HL adapter			
56526486	SRD5: 1:1 fragments			
33109954	SRD5: long-RXN			
56994075	SRD9:1 (CS/SC)	Cancer Cell line	Cancer	
263977	TSC Screen 1			

47

95
 90
 85
 His Ser Gly Gln Asn Leu Tyr Gln Cys
 105
 100
 95
 210 > 5511
 211 > 379
 212 > DNA
 213 > Homo sapiens
 400 > 5511
 tccggagtggt caccagccttc agccacaagg ctttcctgat tgggctccac atctgcagaa
 60
 ccttccttgg gaaaagagg catcgttcca atcgcatagt cacacacatc ccttaactca
 120
 ctctgctgag tctgtgtag tctgtgttcc tcttccact tataggatgg gtccctcatc
 180
 tcttagactt caagcccca ggcagagacc tggctgtccc tcatggagac ctcaaggatata
 240
 atgctgaatt cctctatggc agagatggga ggaagagctc cagctgggc ctctcagcc
 300
 tccatcaggc ctgaatcctg gtctgtgtca catgtgtctt cggccccaag gtccctccca
 360
 gggtccggcg ccggccgcn
 379
 210 > 5512
 211 > 101
 212 > PRT
 213 > Homo sapiens
 400 > 5512
 Met Gln Ala Gln Arg Gly Ala Ser Pro Ile Ser Ala
 1
 Ile Gln Gln Phe Ser Ile Ile pro Gln Ala Pro Met Arg Ser Ser Gln
 20
 Val Ser Ala Leu Gly Leu Gln Ala Gln Asp Gln Asp Pro Ser Tyr
 35
 Lys Trp Arg Gln Gln His Arg Leu Ser Ala Thr Gln Ser Gln Leu
 50
 Arg Asp Val Cys Asp Tyr Ala Ile Gln Thr Met Pro Ser Phe pro Lys
 65
 Gln Gly Ser Ala Asp Val Gln pro Asn Gln Ser Leu Val Ala Gln
 85
 Ala Cys Asp Thr Pro
 100
 210 > 5513
 211 > 837
 212 > DNA
 213 > Homo sapiens
 400 > 5513
 nnaagcttga gtctctctgt ccaagggccag ggaacctgtga ccatgggaagc agagagaagc
 60

aagggcagcg cggcgccctt ggagagtttc cggcaggttg gcccggcgga gctgtcgtg
 agatcggggg agccattgac catcgtctct gaggatggag actggtggag ggtgctgtct
 gaagtctcag gcagagagta taacatcccc aggttccacg tggccaagt cttccatggg
 240
 tggctgtatg agggcctgag caggagagaa gcagagagacc tgggtgtgtt acctgggaa
 300
 cctggagggg ccttcctcat cgggagagag cagaccagga gaggtcttta cttctgtca
 360
 gtcggcctca gcccgcctgc atcctgggag cggatcagac actcagagat ccactgctt
 420
 gacaatggct ggtgtatcat ctcaaccgag ctcaacttcc cttcaatcca gggcctgtg
 480
 gaccattact ctgagctggc gtagatgacac tgcgtgctac tcaaggagcc ctgtgtctg
 540
 cagagggtg gcccgcctcc tggcagagat ataccctac ctgtgactgt gcagagagaca
 600
 ccaatcaact ggaagagct ggaagctcc ctccgtttt ctgaagctgc cacaggggag
 660
 gactcttc tcagtgggg tctcgggag tccctcagct tctacatcag cctgaatgac
 720
 gaggtctgtc ctttgatga tgcctagggc caaaggagag gccaaaagg aaaccaaggc
 780
 tgcacacta gaaccat tcaagcttcc tggcaccga gaggaaggc tgtgac
 837

<210> 5514

<211> 248

<212> PRT

<213> Homo sapiens

<400> 5514

Xaa Ser Leu Ser Ser Val Gln Gly Pro Val Thr Met Gln
 1
 Ala Gln Arg Ser Lys Ala Thr Ala Ala Leu Gly Ser Phe Pro Ala
 20
 Gly Gly Pro Ala Gln Leu Ser Leu Arg Leu Gly Gln Pro Leu Thr Ile
 35
 Val Ser Gln Asp Gly Asp Trp Trp Thr Val Leu Ser Gln Val Ser Gly
 50
 Arg Gln Tyr Asn Ile Pro Ser Val His Val Ala Lys Val Ser His Gly
 65
 Trp Leu Tyr Gln Gly Leu Ser Arg Gln Lys Ala Gln Asp Leu Leu Leu
 85
 Leu Pro Gly Asn Pro Gly Gly Ala Phe Leu Ile Arg Gln Ser Gln Thr
 100
 Arg Arg Gly Ser Tyr Ser Leu Ser Val Arg Leu Ser Arg Pro Ala Ser
 115
 Trp Asp Arg Ile Arg His Tyr Arg Ile His Cys Leu Asp Asn Gly Trp
 130
 Leu Tyr Ile Ser Pro Arg Leu Thr Phe Pro Ser Leu Gln Ala Leu Val
 145
 Asp His Tyr Ser Gln Leu Ala Asp Asp Ile Cys Cys Leu Leu Lys Gln
 160

165
Pro Cys Val Leu Arg Ala Gly Pro Leu Pro Gly Lys Asp Ile Pro
170
180
185
190
195
200
205
210
215
220
225
230
235
240
245
Glu Ala Val Ser Leu Asp Ala

<210> 5515
<211> 420
<212> DNA
<213> Homo sapiens

<400> 5515
gtttgtacca accctctc catccttgaa gcagtcaggg ccaactgcaa gaaatgcaa
60
gaaagatgt ccgcacagct gctgctgct gagagcagac aaaaagaact ggaatggag
120
aagctcagc taccaagccct tgagcaagag caaagaagc tggtgcccgg ccttgaggaa
180
gagctggca aagaacagca ggtggtcctg atgctggtca aagatgcaa gcagctctca
240
agcaagtca tagagggagc ccagaagctc gaagacgtaa tggccaact ggtctctct
300
ccttgacc agcaactgt tcatagctc tctgagtgcc caggaaagg tcatatagat
360
taaatccc ataccgtccc tggataata cctcctcct gcgagcccg agggcctcga
420

<210> 5516
<211> 120
<212> PRT
<213> Homo sapiens

<400> 5516
Val Cys Thr Asn Pro Leu Ser Ile Leu Glu Ala Val Met Ala His Cys
1
Lys Lys Met Glu Arg Met Ser Ala Glu Leu Ala Ala Glu Ser
20
25
30
35
40
45
50
55
60
65
70
75
80
85
90
95
100
105
110
Leu Ala Ser Ser Leu Cys His Glu His Leu Leu His Ser Leu Ser Gly
Val Pro Gly Thr Gly His Ile Asp